From:	Bowman, Amy
Sent:	Tuesday, June 28, 2005 1:39 PM
To:	STIC-Biotech/ChemLib
Cc:	Bowman, Amy
Subject:	sequence search-10/735,991

Hello,
I need SEQ ID NO: 5 in application 10/735,991 searched.
Thank you,
Amy Bowman
AU 1635
REM 2C31
571-272-0755   *1014NA*

Beverly shears

Date completed: _____	**Search Site**	**Vendors**
Searcher: _Beverly e 2528_	_____ STIC	_____ IG
Terminal time: _____	_____ CM-1	_____ STN
Elapsed time: _____	_____ Pre-S	_____ Dialog
CPU time: _____	**Type of Search**	_____ APS
Total time: _____	_____ N.A. Sequence	_____ Geninfo
Number of Searches: _____	_____ A.A. Sequence	_____ SDC
Number of Databases: _____	_____ Structure	_____ DARC/Questel
	_____ Bibliographic	__✓__ Other _CGN_

This Page Blank (uspto)

(1)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: July 5, 2005, 08:57:44 ; Search time 645 Seconds
                                       (without alignments)
                                       9306.384 Million cell updates/sec

Title: US-10-735-991-5
Perfect score: 1014
Sequence: 1 atgaactcgtgggacgcgggg.........ttctgccggtgtctgagtga 1014

Scoring table: IDENTITY_NUC
Searched: Gapop 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 8780412
Searched: 4390206 seqs, 2959870667 residues

Post-processing: Minimum Match 0%
                 Maximum Match 100%
Minimum DB seq length: 0   Listing first 45 summaries
Maximum DB seq length: 200000000

Database :   N_Geneseq_16Dec04:*
                 1: geneseqn1980s:*
                 2: geneseqn1990s:*
                 3: geneseqn2000s:*
                 4: geneseqn2001as:*
                 5: geneseqn2001bs:*
                 6: geneseqn2002as:*
                 7: geneseqn2002bs:*
                 8: geneseqn2003as:*
                 9: geneseqn2003bs:*
                 10: geneseqn2003cs:*
                 11: geneseqn2003ds:*
                 12: geneseqn2004as:*
                 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1014	100.0	1014	12	ADQ94493	Adq94493 Human G-p
2	1011	99.7	1011	12	ADQ94489	Adq94489 Mouse G-p
3	1010.8	99.7	1014	6	ABK47595	Abk47595 cDNA enco
4	1010.8	99.7	7512	12	ADO29763	Ado29763 Human nov
5	1009.2	99.5	2530	13	ADQ89177	Adq89177 Human uro
6	1009.2	99.5	2581	2	AAZ00034	Aaz00034 G protein
7	1006.2	99.2	1728	10	ADF70549	Adf70549 Orphan re
8	986.0	97.3	1014	2	ABS73406	Abs73406 cDNA enco
9	869.2	85.7	1011	12	ADQ29057	Ado29057 Mouse nov
10	868.4	85.6	1172	12	ADQ94491	Adq94491 Rat G-pro
11	556.6	54.9	565	9	AAS68630	Aas68630 DNA encod
12	472.2	46.6	585	9	ACH39042	Ach39042 Human foe
13	418.2	41.2	1092	11	AAS07950	Aas07950 Human cDN
14	418.2	41.2	1092	11	ADI96477	Adi96477 Human G p
15	416.6	41.1	5665	10	AB224430	Ab224430 Human con
16	415	40.9	1092	3	AAC64108	Aac64108 Human G p
17	415	40.9	1092	4	AAF59997	Aaf59997 Human G p
18	415	40.9	1092	4	AAD08865	Aad08865 Human G-p
19	415	40.9	1092	4	AAH73502	Aah73502 Human G p
20	415	40.9	1092	5	AAS08268	Aas08268 Human cDN
21	415	40.9	1092	6	AAD26313	Aad26313 Human G-p
22	415	40.9	1092	8	ABA00881	Aba00881 Somatosta
23	415	40.9	1092	11	ADL96541	Adl96541 Human mut
24	415	40.9	1092	13	ADS09774	Ads09774 Human the
25	415	40.9	1163	6	ABL41643	Abl41643 Human G-p
26	415	40.9	1188	4	AAH51105	Aah51105 Human nGP
27	415	40.9	1352	6	ABS70338	Abs70338 DNA encod
28	415	40.9	1510	4	ABK53077	Abk53077 cDNA enco
29	415	40.9	1510	4	AAS46219	Aas46219 Human DNA
30	415	40.9	1510	8	ACA89669	Aca89669 cDNA enco
31	415	40.9	1510	8	ACA73679	Aca73679 Human sec
32	415	40.9	1510	8	ACA05994	Aca05994 Human sec
33	415	40.9	1510	8	ACA66828	Aca66828 Human sec
34	415	40.9	1510	8	ACF20403	Acf20403 Human sec
35	415	40.9	1510	8	ACF19789	Acf19789 Human sec
36	415	40.9	1510	8	ACD22077	Acd22077 Human sec
37	415	40.9	1510	8	ACF13242	Acf13242 Human sec
38	415	40.9	1510	8	ACD25345	Acd25345 Human sec
39	415	40.9	1510	8	ACF00394	Acf00394 Human sec
40	415	40.9	1510	8	ACA72451	Aca72451 Novel hum
41	415	40.9	1510	8	ACD04975	Acd04975 Novel hum
42	415	40.9	1510	8	ACD18436	Acd18436 Novel hum
43	415	40.9	1510	8	ACD08443	Acd08443 Human sec
44	415	40.9	1510	8	ACA88877	Aca88877 Novel hum
45	415	40.9	1510	8	ACA70319	Aca70319 Human sec

ALIGNMENTS

RESULT 1
ADQ94493

ID   ADQ94493 standard; cDNA; 1014 BP.
XX
AC   ADQ94493;
XX
DT   23-SEP-2004 (first entry)
XX
DE   Human G-protein coupled receptor sequence 115 cDNA.
XX
KW   anorectic; immunomodulator; gene therapy; body weight;
KW   mammalian sequence 115; obesity; cachexia; G protein coupled receptor;
KW   sequence 115; human; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   CDS   1..1014
FT   /tag= a
FT   /product= "Human G-protein coupled receptor sequence 115"
XX
PN   US2004121395-A1.
XX
PD   24-JUN-2004.
XX
PF   15-DEC-2003; 2003US-00755991.
XX
PR   23-DEC-2002; 2002US-0436375P.
XX
PA   (GOOD/) GOODNOW R A.
PA   (MARK/) MARK D F.
PA   (MART/) MARTIN M L.
PA   (ROSI/) ROSINSKI J A.
XX
PI   Goodnow RA, Mark DF, Martin ML, Rosinski JA;
XX
DR   WPI; 2004-468185/44.
DR   N-PSDB; ADQ94494.
XX
PT   Identifying compounds, useful for modulating body weight and for treating
PT   obesity and cachexia, comprises contacting a test compound with a
PT   mammalian sequence 115.

```
Claim 36; SEQ ID NO 5; 28pp; English.

The invention describes a method of identifying compounds useful for
modulating body weight, comprising contacting a test compound with a
mammalian sequence 115. The method comprises: contacting the test compound
with a mammalian sequence 115; determining whether the test compound
binds to the mammalian sequence 115; and identifying a compound that
binds to the mammalian sequence 115 as a compound useful for modulating
body weight. Also described are: a pharmaceutical formulation or
composition, for modulating body weight, comprising a compound that
modulates the activity of a mammalian sequence 115, mixed with a
pharmaceutical carrier; a package comprising the pharmaceutical
formulation and instructions for administering the pharmaceutical
formulation for modulating body weight; preparing a pharmaceutical
composition useful for modulating body weight; treating obesity and
cachexia; and an antibody that recognises an isolated polypeptide
comprising a sequence of 1014 or 337 amino acids (SEQ ID NOS: 5 or 6).
The method is useful for identifying compounds useful for modulating body
weight. The formulation or composition and methods are useful for
treating obesity and cachexia. This sequence encodes human G protein
coupled receptor sequence 115.
```

SQ Sequence 1014 BP; 152 A; 363 C; 301 G; 198 T; 0 U; 0 Other;

```
Query Match    100.0%;  Score 1014;  DB 12;  Length 1014;
Best Local Similarity  100.0%;  Pred. No. 1.2e-172;
Matches 1014;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    1 ATGAACTCGTGTGGGACGCGGCCGCCTGGCGGGCGTACTGTGGTGGGCCACCATGGGCGGTCTCGGCTG 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 ATGAACTCGTGTGGGACGCGGCCGCCTGGCGGGCGTACTGTGGTGGGCCACCATGGGCGGTCTCGGCTG 60

Qy   61 CTGTCCAACCGGCGTCGGTCGTTGCCTCGTGCACAGAGGCGGACTCCCCGCCCAGGCG 120
Db   61 CTGTCCAACCGGCGTCGGTCGTTGCCTCGTGCACAGAGGCGGACTCCCCGCCCAGGCG 120

Qy  121 CCGGGCGCTCTTCACCCTGAACCTCAGGTGCGGGAACCTGCTGTTGCACGTTGGTCAACATG 180
Db  121 CCGGGCGCTCTTCACCCTGAACCTCAGGTGCGGGAACCTGCTGTTGCACGTTGGTCAACATG 180

Qy  181 CCGCCTCACGCTGGCCGTGGTGGGCCAGCGGCCACCGCCCTGTGCCGC 240
Db  181 CCGCCTCACGCTGGCCGTGGTGGGCCAGCGGCCACCGCCCTGTGCCGC 240

Qy  241 CTGGGCTGCCTTCCTCGGACACCTTCTCCATTGCCTCAGCATCGAGGCCCGGCTC 300
Db  241 CTGGGCTGCCTTCCTCGGACACCTTCTCCATTGCCTCAGCATCGAGGCCCGGCTC 300

Qy  301 AGCATCGACCGGCCGTGGGGTTCCGGCTGCTACCGGGGCCAAGATGCGCCTC 360
Db  301 AGCATCGACCGGCCGTGGGGTTCCGGCTGCTACCGGGGCCAAGATGCGCCTC 360

Qy  361 CGCGACGGCGGGCCTCAGTGGTGGCTACACGTGGCTGCACGGGCTCCCAGCCGCC 420
Db  361 CGCGACGGCGGGCCTCAGTGGTGGCTACACGTGGCTGCACGGGCTCCCAGCCGCC 420

Qy  421 GCGCTCGCCCGTGTCTCGGGCTTCCACCAGCGTGACCGGCCTGCAGGCTGCAGC 480
Db  421 GCGCTCGCCCGTGTCTCGGGCTTCCACCAGCGTGACCGGCCTGCAGGCTGCAGC 480

Qy  481 CGGGCGGCCAGACGAGGAGCCCTCGCGGTTCAGCCTTCCAGGCTCTCAGC 540
Db  481 CGGGCGGCCAGACGAGGAGCCCTCGCGGTTCAGCCTTCCAGGCTCTCAGC 540

Qy  541 TTCCTGCTCTCCTTCGTGCGTGCTGCACGTAGCTCAAGGTGTGGGCCGC 600
Db  541 TTCCTGCTCTCCTTCGTGCGTGCTGCACGTAGCTCAAGGTGTGGGCCGC 600

Qy  601 TTCCATTGCAAGCGACGTGATCACCAGTGCGTGGTGCTGGTGACCTG 660
Db  601 TTCCATTGCAAGCGACGTGATCACCAGTGCGTGGTGCTGGTGACCTG 660

Qy  661 CACCCCAGTGTGCGGGAACGCTGTCTGGAGGAGCAGGAGCAGGAGGCCACC 720
Db  661 CACCCCAGTGTGCGGGAACGCTGTCTGGAGGAGCAGGAGCAGGAGGCCACC 720
```

RESULT 2
ADQ94489
ID   ADQ94489 standard; cDNA; 1011 BP.
XX
AC   ADQ94489;
XX
DT   23-SEP-2004 (first entry)
XX
DE   Mouse G-protein coupled receptor sequence 115 cDNA.
XX
KW   anorectic; immunomodulator; gene therapy; body weight;
KW   mammalian sequence 115; obesity; cachexia; G protein coupled receptor;
KW   sequence 115; mouse; gene; ss.
XX
OS   Mus musculus.
XX
FH
FT   Key   Location/Qualifiers
FT   CDS   1..1011
FT   /tag= a
FT   /product= "Mouse G-protein coupled receptor sequence 115"
XX
PN   US2004121395-A1.
XX
PD   24-JUN-2004.
XX
PF   15-DEC-2003; 2003US-00735991.
XX
PR   23-DEC-2002; 2002US-0436375P.
XX
PA   (GOOD/) GOODNOW R A.
PA   (MARK/) MARK D F.
PA   (MART/) MARTIN M L.
PA   (ROSI/) ROSINSKI J A.
XX
PI   Goodnow RA, Mark DF, Martin ML, Rosinski JA;
XX
DR   WPI; 2004-468185/44.
DR   P-PSDB; ADQ94490.
XX
PT   Identifying compounds, useful for modulating body weight and for treating
PT   obesity and cachexia, comprises contacting a test compound with a
PT   mammalian sequence 115.
XX
PS   Example 3; SEQ ID NO 1; 28pp; English.

The invention describes a method of identifying compounds useful for
modulating body weight, comprising contacting a test compound with a
mammalian sequence 115. The method comprises: contacting a test compound
with a mammalian sequence 115; determining whether the test compound

```
Db  661 CACCCCAGTGTGCGGGAACGCTGTCTGGAGGAGCAGGAGCAGGAGGCCACC 720

Qy  721 AAGAAGGATCAGCAACCTTCATTAGGAGGACCTTCATTCGGCGGCCCTTATGTGGATCACC 780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721 AAGAAGGATCAGCAACCTTCATTAGGAGGACCTTCATTCGGCGGCCCTTATGTGGATCACC 780

Qy  781 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGGCTGGTGCTTCGGCTTCGGCTTGCTTCGG 840
Db  781 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGGCTGGTGCTTCGGCTTCGGCTTGCTTCGG 840

Qy  841 TGCTTGGCGTACAGCAAGGGCCGCATCGGACCCCCTTTGTGTACTCCTTACTGGCGACACCAG 900
Db  841 TGCTTGGCGTACAGCAAGGGCCGCATCGGACCCCCTTTGTGTACTCCTTACTGGCGACACCAG 900

Qy  901 TACCGGCCAAAAGCTGCAAGGAGATTCTGAACAGGAGGCTCCTGGACAGCACGCGCTCCACTCC 960
Db  901 TACCGGCCAAAAGCTGCAAGGAGATTCTGAACAGGAGGCTCCTGGACAGCACGCGCTCCACTCC 960

Qy  961 TCTGGCCTCACAGGCGGCGACTCTCACAGCCAGAGAACAATTCTGCCGGTGTCTGAGTGA 1014
Db  961 TCTGGCCTCACAGGCGGCGACTCTCACAGCCAGAGAACAATTCTGCCGGTGTCTGAGTGA 1014
```

```
CC   binds to the mammalian sequence 115; and identifying a compound that
CC   binds to the mammalian sequence 115 as a compound useful for modulating
CC   body weight. Also described are: a pharmaceutical formulation or
CC   composition, for modulating body weight, comprising a compound that
CC   modulates the activity of a mammalian sequence 115, mixed with a
CC   pharmaceutical carrier; a package comprising the pharmaceutical
CC   formulation and instructions for administering the pharmaceutical
CC   formulation for modulating body weight; preparing a pharmaceutical
CC   composition useful for modulating body weight; treating obesity and
CC   cachexia; and an antibody that recognises an isolated polypeptide
CC   comprising a sequence of 1014 or 337 amino acids (SEQ ID NOS: 5 or 6).
CC   The method is useful for identifying compounds useful for modulating body
CC   weight. The formulation or composition and methods are useful for
CC   treating obesity and cachexia. This sequence encodes G protein coupled
CC   receptor sequence 115.
XX
SQ   Sequence 1011 BP; 151 A; 363 C; 300 G; 197 T; 0 U; 0 Other;

Query Match             99.7%;  Score 1011;  DB 12;  Length 1011;
Best Local Similarity   100.0%;  Pred. No. 4e-172;
Matches 1011; Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 ATGAACTGTGGGACTGGGACGGCGGGGGCCTGGCCGGGCCGGGGCCTTACTGGTGGGGCAGCGATGGGGCGGGTCTCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGAACTGTGGGACTGGGACGGCGGGGGCCTGGCCGGGCCGGGGCCTTACTGGTGGGGCAGCGATGGGGCGGGTCTCGCTG 60

Qy         61 CTGTCCAACCGGCTGGTGCTCTCTGCCTCGCTTGCGCTGCACAGCGCGGGACATCCCGCGCCCAGGCG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGTCCAACCGGCTGGTGCTCTCTGCCTCGCTTGCGCTGCACAGCGCGGGACATCCCGCGCCCAGGCG 120

Qy        121 CCGGGGGCTCTTCACCCTGAACCTCAGTGCGCGGGAACCTGCGTGGTCAACGTGGTCAACATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGGGGGCTCTTCACCCTGAACCTCAGTGCGCGGGAACCTGCGTGGTCAACGTGGTCAACATG 180

Qy        181 CCGGCTCACGCTGGCCGCGTGTGGGCGCAGCGGGCCACCGCGCCCTGTCCGCGC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCGGCTCACGCTGGCCGCGTGTGGGCGCAGCGGGCCACCGCGCCCTGTCCGCGC 240

Qy        241 CTGGCTGCCTTCCTCGACACACTTCCTGGGCTCCAACTCCANTGCCATGGCCGGCTC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGCTGCCTTCCTCGACACACTTCCTGGGCTCCAACTCCANTGCCATGGCCGGCTC 300

Qy        301 AGCATCGACCGCGTGGCCGGTGGTCTTCCGGCTACCGGGCCAAGATGCGCCTC 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCATCGACCGCGTGGCCGGTGGTCTTCCGGCTACCGGGCCAAGATGCGCCTC 360

Qy        361 CGCGACGCGCGGCCGCTCATGGTGGCCTACACGTGGCTGCCGTGGCCTCACGCGCC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGACGCGCGGCCGCTCATGGTGGCCTACACGTGGCTGCCGTGGCCTCACGCGCC 420

Qy        421 GCGGCTGCGCCCGTGCTCGTGGCGTTGCGGCTTGTACCGCTCGCTGCGACGCTGCAGC 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCGGCTGCGCCCGTGCTCGTGGCGTTGCGGCTTGTACCGCTCGCTGCGACGCTGCAGC 480

Qy        481 CGGCGGCCCAGACGAGCGCCCTGGCGTTCGCCGTGCCGCCTTCCACGCTCTCAGC 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGGCGGCCCAGACGAGCGCCCTGGCGTTCGCCGTGCCGCCTTCCACGCTCTCAGC 540

Qy        541 TTCCTGCTCTCCTTGGTCGTGCGCAGTACCTCAAGGTGCTCAAGGTGGCCCGC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCTGCTCTCCTTGGTCGTGCGCAGTACCTCAAGGTGCTCAAGGTGGCCCGC 600

Qy        601 TTCCATTGCAAGCCCATCGACGTGATCACCATGCGAGACGCTGGTGGTGCTGGACCTG 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCATTGCAAGCCCATCGACGTGATCACCATGCGAGACGCTGGTGGTGCTGGACCTG 660

Qy        661 CACCCCAGTGTGCGGGAACGGCTGTGGAGAGCAGAGAGCGGAGGCGACAGCGACCACC 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACCCCAGTGTGCGGGAACGGCTGTGGAGAGCAGAGAGCGGAGGCGACAGCGACCACC 720

Qy        721 AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGTGGCTTCGCGCCCTANGTGTGATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGTGGCTTCGCGCCCTANGTGTGATCACC 780

Qy        781 AGGCTAGTGGGAGCTTCTTCTTCCACGGTGCCCATCGGCTCCCACTGGGGCGGTGCGTGTCCAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGCTAGTGGAGCTTCTTCTTCCACGGTGCCCATCGGCTCCCACTGGGGCGGTGCGTGTCCAAG 840

Qy        841 TGCTTGCGTACAGCAAGGCCGCCATCCGGACCGGCCCCTTTGTGTACTGCCCTTACTGGCACACCAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCTTGCGTACAGCAAGGCCGCCATCCGGACCGGCCCCTTTGTGTACTGCCCTTACTGGCACACCAG 900

Qy        901 TACCGCAAAAGCTGCAAGGAGATTCTGAACAGGGCTTCCACGAGCGGCTTCCATCCACTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACCGCAAAAGCTGCAAGGAGATTCTGAACACAGGGCTTCCACGAGCGGCTTCCATCCACTCC 960

Qy        961 TCTGGGCCTCACAGGCGGACTCTCACAGCCCAGAGACATTCTGCCCGGTGTCTGAG 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTGGGCCTCACAGGCGGACTCTCACAGCCCAGAGACATTCTGCCCGGTGTCTGAG 1011
```

RESULT 3
ABK47595
ID   ABK47595 standard; cDNA; 1014 BP.
XX
AC   ABK47595;
XX
DT   18-JUN-2002 (first entry)
XX
DE   cDNA encoding human G protein coupled receptor AXOR49 protein.
XX
KW   Human; G protein coupled receptor AXOR49; receptor; bacterial infection;
KW   viral infection; fungal infection; protozoal infection; pain; cancer;
KW   tumour; HIV; human immunodeficiency virus; diabetes; Parkinson's disease;
KW   heart failure; allergy; vaccine; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1014
FT   /tag= a
FT   /product= "Human G protein coupled receptor AXOR49"
XX
PN   GB2364057-A.
XX
PD   16-JAN-2002.
XX
PF   03-APR-2001; 2001GB-00008242.
XX
PR   04-APR-2000; 2000US-00542981.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
PA   (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI   Elshourbagy N;
XX
DR   WPI; 2002-306798/35.
DR   P-PSDB; AAU78644.
XX
PT   New G protein-coupled receptor AXOR49, useful for diagnosis and treatment
PT   of e.g. infections and cancer.
XX
PS   Claim 2; Page 27; 34pp; English.
XX
CC   The present invention relates to a new polypeptide encoded by a
CC   polynucleotide comprising a fully defined sequence of 1014 base pairs, as
CC   given in the specification. The polypeptide of the invention comprises a
CC   sequence at least 95% identical to a fully defined sequence of 337 amino
CC   acids as given in the specification, or a fragment or variant. The
CC   polypeptide, a putative G protein-coupled receptor designated AXOR49, and
CC   the nucleic acid that encodes it are useful for treating a very wide
CC   range of diseases, e.g. infections (bacterial, viral, fungal or
CC   protozoal, especially human immune deficiency virus-1 or -2), pain,
CC   cancer, diabetes, Parkinson's disease, heart failure and allergy. The
CC   molecules of the invention can also be used to diagnose (susceptibility
CC   to) diseases (including detecting mutations or measuring mRNA or protein
```

```
CC     levels) associated with inappropriate activity or levels of AXOR49, to
CC     screen for (ant)agonists, potential therapeutic agents and in vaccines.
CC     The present nucleic acid sequence encodes the human G protein coupled
CC     receptor AXOR49 protein of the invention
XX
SQ     Sequence 1014 BP; 151 A; 364 C; 301 G; 198 T; 0 U; 0 Other;

Query Match      99.7%;  Score 1010.8;  DB 6;  Length 1014;
Best Local Similarity   99.8%;  Pred. No.4.4e-172;
Matches 1012;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy        1 ATGAACTCGTGGGACGCGGGAGGCGGCGGCGGGCGGCGGCGGCTACTCGTCGCACGAGATGGCGGTCTCGCTG 60
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGAACTCGTGGGACGCGGGAGGCGGCGGCGGGCGGCGGCGGCTACTCGTCGCACGAGATGGCGGTCTCGCTG 60

Qy       61 CTGTCCCAACGCGCGCTGGTGCTGCTGCTGCTGCACAGCCGGACATCCCCGCCCCAGGCG 120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTGTCCCAACGCGCGCTGGTGCTGCTGCTGCTGCACAGCCGGACATCCCCGCCCCAGGCG 120

Qy      121 CCGGCGGCGCTCTTCACCCTGAACCTGCACGTGGGGAACCTGCTGGTCCCCGCCGTCAACATG 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CCGGCGGCGCTCTTCACCCTGAACCTGCACGTGGGGAACCTGCTGGTCCCCGCCGTCAACATG 180

Qy      181 CCCGGCTCACCGTCGGCCCGCGGGCGCAGCGGGACGCGCGGCCCTGTGCCGC 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CCCGGCTCACCGTCGGCCCGCGGGCGCAGCGGGACGCGCGGCCCTGTGCCGC 240

Qy      241 CTGGGCTGCCTTCCTGCGACGACGCCGGGAGCAGCCGGACCCCGTGTGCCGCTC 300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 CTGGGCTGCCTTCCTGCGACGACGCCGGGAGCAGCCGGACCCCGTGTGCCGCTC 300

Qy      301 AGCACATCGACACCGCTGGGTGGGCGGTGGTCTTCCCGCGTACACGGGCCAAGATGCCCTC 360
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 AGCACATCGACACCGCTGGGTGGGCGGTGGTCTTCCCGCGTACACGGGCCAAGATGCCCTC 360

Qy      361 CGCGACGCGGCGGCGCTCATGTGGCGCGTGCTGCTGCACACTTCCCAGCCGCC 420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CGCGACGCGGCGGCGCTCATGTGGCGCGTGCTGCTGCACACTTCCCAGCCGCC 420

Qy      421 GCGCTCGCGCCTGTTCCTGGCCTTCCACCAGAGTGTACGGCCTGCACGTGCCAG 480
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GCGCTCGCGCCTGTTCCTGGCCTTCCACCAGAGTGTACGGCCTGCACGTGCCAG 480

Qy      481 CGGGCGGCCCAGACGACGGCCTTCGCCCTTCACTGGCCACGCCTCTCAGC 540
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 CGGGCGGCCCAGACGACGGCCTTCGCCCTTCACTGGCCACGCCTCTCAGC 540

Qy      541 TTCCTGGCTCTTCCTTCGTTCGTTCGTGCTCTTCGCTCCAAGGTGCTCAAGGTGCCCGC 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 TTCCTGGCTCTTCCTTCGTTCGTTCGTGCTCTTCGCTCCAAGGTGCTCAAGGTGCCCGC 600

Qy      601 TTCCAATTGCAAGCGCAATGACAGTGATCACATGACAAGCGTGGGTGGTCTGGGGTGACCTG 660
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 TTCCAATTGCAAGCGCAATGACAGTGATCACATGACAAGCGTGGGTGGTCTGGGGTGACCTG 660

Qy      661 CACCCCAGTGTGCGGGAACCGCTGTCTGGAGAGCACAGAGCGAGCAGCGAGCCACC 720
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 CACCCCAGTGTGCGGGAACCGCTGTCTGGAGAGCACAGAGCGAGCAGCGAGCCACC 720

Qy      721 AAGAAGATCAGCACACTTCATAGGGACCTTCCTTGTGTGTTCGCGCCTATTGTTGAATCACC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 AAGAAGATCAGCACACTTCATAGGGACCTTCCTTGTGTGTTCGCGCCTATTGTTGAATCACC 780

Qy      781 AGGCTAGTGGAGCTCTTCCTCCACGGTGCCCATGGGGGTGCGTGTCCAAG 840
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 AGGCTAGTGGAGCTCTTCCTCCACGGTGCCCATGGGGGTGCGTGTCCAAG 840

Qy      841 TGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCGTCGACACCAG 900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 TGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCGTCGACACCAG 900

Qy      901 TACCCGCCAAAGCTGCAAGGAGATTCTGAACAGGCTTCGCACGAGACGCTCCATCCACTCC 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 TACCCGCCAAAGCTGCAAGGAGATTCTGAACAGGCTTCGCACGAGACGCTCCATCCACTCC 960

Qy      961 TCTGGCCTCACAGGCCGACTCTCACAGCCAGAAACAATTCTCGCGGTGTCTGAGTGA 1014
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TCTGGCCTCACAGGCCGACTCTCACAGCCAGAACAATTCTCGCGGTGTCTGAGTGA 1014
```

```
RESULT 4
AD029763
ID     AD029763 standard; cDNA; 7512 BP.
XX
AC     AD029763;
XX
DT     29-JUL-2004  (first entry)
XX
DE     Human novel GPCR GPR26 polynucleotide, SEQ ID NO:865.
XX
KW     G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW     transgenic mouse; neurological disorder; adrenal gland disorder;
KW     colon disorder; intestinal disorder; cardiovascular disorder;
KW     muscular disorder; blood disorder; immune disorder; bone disorder;
KW     joint disorder; metabolic disorder; nutritive disorder; cancer;
KW     kidney disorder; liver disorder; lung disorder; breast disorder;
KW     ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW     thymus disorder; stomach disorder; pancreas disorder; spleen disorder;
KW     skin disorder; thyroid disorder; antiparkinsonian; antimanic;
KW     cytostatic; antiinflammatory; antiparkinsonian; antimanic;
KW     CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW     virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW     dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW     immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW     gene; ss.
XX
OS     Homo sapiens.
XX
PN     WO2004040000-A2.
XX
PD     13-MAY-2004.
XX
PF     09-SEP-2003; 2003WO-US028226.
XX
PR     09-SEP-2002; 2002US-0409303P.
PR     09-APR-2003; 2003US-0461329P.
XX
PA     (PRIM-) PRIMAL INC.
XX
PI     Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
PI     Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
XX
DR     WPI; 2004-390329/36.
DR     P-PSDB; AD029055.
XX
PT     Novel mammalian G protein coupled receptors, useful for identifying
PT     compounds that modulates diagnosing and treating disease condition
PT     associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT     pectoris, Parkinson's disease.
XX
PS     Claim 13; SEQ ID NO 865; 542pp; English.
XX
CC     The invention relates to human and mouse G protein-coupled receptors
CC     (GPCRs) and nucleic acids encoding them. The invention also relates to
CC     sequences at least 90% identical to the GPCR proteins and nucleic acids
CC     of the invention; methods of treating, preventing or diagnosing diseases
CC     associated with GPCRs of the invention; methods of screening for
CC     compounds useful in the treatment of GPCR-related diseases; a transgenic
CC     mouse comprising a GPCR gene of the invention; a mouse comprising a
CC     mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC     from the transgenic mice; kits comprising several mice, each of which has
CC     a mutation in a different GPCR gene of the invention; and kits comprising
CC     probes which hybridise to GPCR polynucleotides of the invention. The
CC     invention further discloses variants of the GPCR polypeptides and vectors
CC     comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC     be used in the diagnosis, treatment or prevention of a wide variety of
```

```
Db      901 TACCGCCAAAGCTGCAAGGAGGATTCTGAACAGGCTTCGCACGAGACGGCTCCATCCACTCC 960
              ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      961 TCTGGCCTCACAGGCCGACTCTCACAGCCAGAAACAATTCTCGCGGTGTCTGAGTGA 1014
              |||||||||||||||||||||||||||| |||||||||||||||||||
Db      961 TCTGGCCTCACAGGCCGACTCTCACAGCCAGAACAATTCTCGCGGTGTCTGAGTGA 1014
```

CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 7512 BP; 1779 A; 1828 C; 1976 G; 1929 T; 0 U; 0 Other;

Query Match 99.7%; Score 1010.8; DB 12; Length 7512;
Best Local Similarity 99.8%; Pred. No. 4.6e-172;
Matches 1012; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 ATGAACTCGTGGGGACGCGGGCGGGCCGCCCTGGCCGGGCGGGCGGGGGGGCCACGGCCTGGGCTTCGGCTG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   53 ATGAACTCGTGGGGACGCGGGCGGGCCGCCCTGGCCGGGCGGGCGGGGGGGCCACGGCCTGGGCTTCGGCTG 112

Qy   61 CTGTCCAACGCGCTGCTGCTCTCTCCGCTCTGCTGCACAGCGGCGGACATCCCGCGCCAGGCG 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   113 CTGTCCAACGCGCTGCTGCTCTCTCCGCTCTGCTGCACAGCGGCGGACATCCCGCGCCAGGCG 172

Qy   121 CCGGCGCTCTTCACCCTGAACCTCACGTGCGGGAACCTGCTGTTGCACCGTGGTCAACATG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   173 CCGGCGCTCTTCACCCTGAACCTCACGTGCGGGAACCTGCTGTTGCACCGTGGTCAACATG 232

Qy   181 CCGCTCACGCTGTCGTGGGCGCAGCGGACGCGCCCTGTGCCGC 240
           |||||||||||||||||||||||||||||||||||||||||||
Db   233 CCGCTCACGCTGTCGTGGGCGCAGCGGACGCGCCCTGTGCCGC 292

Qy   241 CTGCTGCCTTCCTCGACACCTTCCATGCCATGAGCGCCGCCTC 300
           |||||||||||||||||||||||||||||||||||||||||||
Db   293 CTGCTGCCTTCCTCGACACCTTCCATGCCATGAGCGCCGCCTC 352

Qy   301 AGCATCGACACCGCTCGGGGTGGGCGGCCGTGGTCTTCCCGGC 360
           |||||||||||||||||||||||||||||||||||||||||||
Db   353 AGCATCGACACCGCTCGGGGTGGGCGGCCGTGGTCTTCCCGGC 412

Qy   361 CGCGACGCGGGGCGCTCATGGTGGCCTCACCTTCCCAGCCGCC 420
           |||||||||||||||||||||||||||||||||||||||||||
Db   413 CGCGACGCGGGGCGCTCATGGTGGCCTCACCTTCCCAGCCGCC 472

Qy   421 GCGCTCGGCCCCTGTCCTGGGCTTACGCCCTCGTGCACGCTGTGCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   473 GCGCTCGGCCCCTGTCCTGGGCTTACGCCCTCGTGCACGCTGTGCAGC 532

Qy   481 CGGCGGGCCAGCACGAGCGAGCGCTTCGCGCGCTTCGCCTTCACGCTTCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   533 CGGCGGGCCGGACGAGGCGCTTCGCGCGCTTCGCCTTCACGCTTCAGC 592

Qy   541 TTCCTGCTCTCCTTCGTCGTGTCTGCTGCACGTACCTCAAGGTGCCCGC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db   593 TTCCTGCTCTCCTTCGTCGTGTCTGCTGCACGTACCTCAAGGTGCCCGC 652

Qy   601 TTCCATTGCAAGCGCATCGACGTGATCACCATGCAGACGCTGGTTGCTCTGGGTGGACCTG 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   653 TTCCATTGCAAGCGCATCGACGTGATCACCATGCAGACGCTGGTTGCTCTGGGTGGACCTG 712

Qy   661 CACCCCAGTGTGTGCGGGAACGCTGTCTGAGGAGCAGAAGCGGAGGAGGGCCACC 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   713 CACCCCAGTGTGCGGGAAGCTGTCTGGAGGAGCAGAAGCGGAGCAGCGACCACC 772

Qy   721 AAGAAGATCAGCACCTTCATAGGACCTCCTTGTGTGTTCGGGCCCTATGTGATCACC 780

RESULT 5
AC ADQ89177
ID ADQ89177 standard; cDNA; 2530 BP.
XX
AC ADQ89177;
XX
DT 21-OCT-2004 (first entry)
XX
DE Human urological disorder related protein 2882 encoding cDNA SEQ.129.
XX
KW urological disorder; uropathic; cytostatic; urinary incontinence;
KW benign prostatic hyperplasia; human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key   Location/Qualifiers
FT CDS   186..1199
FT   /tag= a
FT   /product= "urological disorder related protein 2882"
XX
PN WO200406557-A2.
XX
PD 05-AUG-2004.
XX
PF 14-JAN-2004; 2004WO-US000750.
XX
PR 15-JAN-2003; 2003US-0440318P.
PR 04-FEB-2003; 2003US-0444783P.
PR 27-MAR-2003; 2003US-0457901P.
PR 08-MAY-2003; 2003US-0468775P.
PR 19-MAY-2003; 2003US-0471614P.
PR 16-JUN-2003; 2003US-0478742P.
PR 18-JUL-2003; 2003US-0488529P.
PR 30-JUL-2003; 2003US-0491156P.
PR 02-SEP-2003; 2003US-0499594P.
PR 26-SEP-2003; 2003US-0506332P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Karichetti V, Silos-Santiago I, Eliasof SD;
XX
DR WPI; 2004-562167/54.
DR N-PSDB; ADQ89178.
XX
PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,
PT 211 or for identifying a compound capable of treating a urological
PT disorder or identifying and treating a subject having a urological
PT disorder.
XX
PS Claim 1; SEQ ID NO 129; 542pp; English.
XX
CC The present invention describes the use of polypeptides related to
CC urological disorders for identifying a compound capable of treating a

---

Db   773 AAGAAGATCAGCACCTTCATAGGGACCTCCTTGTGTGTTCGGGCCGCCCTATGTGATCACC 832
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   781 AGGGCTAGTGGGAGCTCTTCCACGCGTGCCCATCGGCGCTCCCACTGGGCGGTCGTTGCTCCAAG 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   833 AGGGCTAGTGGGAGCTCTTCCACGCGTGCCCATCGGCGCTCCCACTGGGCGGTCGTTGCTCCAAG 892
Qy   841 TGCTTGGCGTACAGCAGGCGCCGCATCCGACCCCTTTGTGTACTCCTTACTGGCACACCAG 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   893 TGCTTGGCGTACAGCAGGCGCCGCATCCGACCCCTTTGTGTACTCCTTACTGGCACACCAG 952
Qy   901 TACCGCAAAAGCTGCAAGGAGATTCTGAACACAGGCTCCTGCACAGAGGCGCTCCATCCACTCC 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   953 TACCGCAAAAGCTGCAAGGAGATTCTGAACACAGGCTCCTGCACAGAGGCGCTCCATCCACTCC 1012
Qy   961 TCTGGGCCTCACAGGCGGCGACTCTCACAGCCAGAGAACATTCTGCCGGGTGTCTGAGTGA 1014
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1013 TCTGGGCCTCACAGGCGGCGACTCTCACAGCCAGAGAACATTCTGCCGGGTGTCTGAGTGA 1066

```
CC   urological disorder, identifying a subject having a urological disorder,
CC   or treating a subject having a urological disorder. Also described: (1) a
CC   method for identifying a compound capable of treating a urological
CC   disorder; (2) a method for identifying a subject having a urological
CC   disorder; and (3) a method for treating a subject having a urological
CC   disorder. The compound has uropathic and cytostatic activities. The
CC   polypeptides related to urological disorders are useful for identifying a
CC   compound capable of treating a urological disorder, identifying a subject
CC   having a urological disorder, or treating a subject having a urological
CC   disorder. Disorders include urinary incontinence and benign prostatic
CC   hyperplasia. The present sequence encodes a human urological disorder
CC   related protein, which is used in the exemplification of the present
CC   invention.
XX
SQ   Sequence 2530 BP; 487 A; 724 C; 739 G; 577 T; 0 U; 3 Other;
```

```
Query Match             99.5%;  Score 1009.2;  DB 13;  Length 2530;
Best Local Similarity   99.7%;  Pred. No. 8.7e-172;
Matches 1011;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy      1 ATGGAACTCGTGGGAGCAGCGGCGGCCGGCGGCCTACTGGTGGGCAGCAGATGGGGGTCTCGCTG     60
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db    186 ATGGAACTCGTGGGACGGCGGCCCTGCGCGGCCGGCCTACTGGTGGGCAGCAGATGGGGGTCTCGCTG    245

Qy     61 CTGTCCAACGGCGCTGGTGTGCTCGTTCGCCTGCTGCACAGCGGGCATCCCGCCCGCCCAGGCG    120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    246 CTGTCCAACGGCGCTGGTGTGCTCGTTCGCCTGCTGCACAGCGGGCATCCCGCCCGCCCAGGCG    305

Qy    121 CCGGCGCGCTTCACCCTGAACCTCACGTGCGGGAACCTGCTTGTGTGCACGTGGTCAACATG    180
          ||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    306 CCGGCGCGCTCTTCACCCTGAACCTCACGTGCGGGAACCTGCTTGTGTGCACGTGGTCAACATG    365

Qy    181 CCGGCCTTCACGGTCTGGCCGCGGCGTGTTGGCGCGGCACCGGGCAGCCGGGCCAAGATGCCCTG    240
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    366 CCGGCTCACGTGGCCGCGGCGTGTTGGCGCGGCACCGGGCAGCCGGGCCAAGATGCCCTG    425

Qy    241 CTGGGCTGCCTTCCTCCTGGACACCTTCCTGGCTGCCGCCAACTCATCCATGCTCAGCATCAGCATGGCCGCCTC    300
          ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db    426 CTGGGCTGCCTTCCTCCTGGCTGCCGCCAACTCATCCATGCTCAGCATCAGCATGGCCGCCTC    485

Qy    301 AGCATCGACCCGCTGGGTGGCGTGGTGCTTCTTCCCGGTACCGGGCCACGCGCCTGTGCCCGC    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    486 AGCATCGACCCGCTGGGTGGCGTGGTGCTTCTTCCCGGTACCGGGCCACGCGCCTGTGCCCGC    545

Qy    361 CGCGACGCGCGGCGGCCTCATGGTGCGGCTACACGTGGCGCGCAGCGGCTCACCTTCCCACCGCC    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 CGCGACGCGCGGCGGCCTCATGGTGCGGCTACACGTGGCGCGCAGCGGCTCACCTTCCCACCGCC    605

Qy    421 GCGCTTCGGCGCCCTGTTCTTGGCGTTCCACCAGCTGTACCGGCCTCGTGTGCACGCTGTGCAGC    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    606 GCGCTTCGGCGCCCTGTTCTTGGCGTTCCACCAGCTGTACCGGCCTCGTGTGCACGCTGTGCAGC    665

Qy    481 CGGGCGCCAGACGAGCGACGGCGCTTCGCGCGTGTTCGCCGCTTCAACTGCGGCTCTCAGC    540
          |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    666 CGGGCGCCGCCGACGAGCGACGGCGCTTCGCGCGTGTTCGCCGCTTCAACTGCGGCTCTCAGC    725

Qy    541 TTCCTGCTCTTCCTTCCTTCGTCGTTCGCACGTACCTGCAGGTGCTCAAGGTGGCCCC    600
          ||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    726 TTCCTGCTCTTCCTTCGTCGTTCGCACGTACCTGCAGGTGCTCAAGGTGGCCCCGC    785

Qy    601 TTCCATTGCAAGCGCATCGACGTGATCACAGTCGTGCTGTGGTGGACCTG    660
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    786 TTCCATTGCAAGCGCATCGACGTGATCACAGTCGTGCTGTGGTGGACCTG    845

Qy    661 CACCCCAGTGTGCGGGAAACCGTGTCTGGAGGAGCAGAGCAGCGAGCCACC    720
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    846 CACCCCAGTGTGCGGGAAACCGTGTCTGGAGGAGCAGAGCAGCGAGCCACC    905

Qy    721 AAGAGAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGTCGCTTCGCGCGTGATCACC    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 AAGAGAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGTCGCTTCGCGCGTTCATCACC    965

Qy    781 AGGCTAGTGGAGCTCTTCCCACGGTGCCCATCGGCTCCCACTGGGGTGGGTGTCCCAAG    840
          ||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    966 AGGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCCCACGGTGCCATCCGGCGCTCCCACTGGGGGGTGCTGTCCAAG   1025

Qy    841 TGCTTGGCGTACAGCAGCAAGGCCGCATCCGGACCCCTTTGTGTACTCCTTACTGCGACAACCAG    900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1026 TGCTTGGCGTACAGCAGCAAGGCCGCATCCGGACCCCTTTGTGTACTCCTTACTGCGACAACCAG   1085

Qy    901 TACCCGCAAAAGCTGCAAAGGAGGATTTCTGAACAGGAGCTCCTGCACCAGAGCTCCATCCACTCC    960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1086 TACCCGCAAAAGCTGCAAAGGAGGATTTCTGAACAGGAGCTCCTGCACCAGAGCTCCATCCACTCC   1145

Qy    961 TCTGCCCTCACAGGGCCGGACTCTCACAGCCCAGAACATTCTGCCGGGTGTCTGAGTGA   1014
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1146 TCTGCCCTCACAGGGCCGGACTCTCACAGCCCAGAACATTCTGCCGGGTGTCTGAGTGA   1199
```

```
RESULT 6
AAZ00034
ID   AAZ00034 standard; cDNA; 2581 BP.
XX
AC   AAZ00034;
XX
DT   12-OCT-1999 (first entry)
XX
DE   G protein coupled receptor flh2882 gene.
XX
KW   G protein coupled receptor; flh2882; brain; signal transduction;
KW   transmembrane domain; ss.
XX
OS   Homo sapiens.
XX
FH   Key              Location/Qualifiers
FT   CDS              184..1197
FT                    /tag= a
FT                    /*tag= a
FT                    /product= "flh2882"
FT                    /note= "G protein coupled receptor"
XX
PN   WO9937679-A1.
XX
PD   29-JUL-1999.
XX
PF   20-JAN-1999;   99WO-US001128.
XX
PR   26-JAN-1998;   98US-00013634.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Glucksmann MA, Robison K;
XX
DR   WPI; 1999-469116/39.
DR   P-PSDB; AAY28488.
XX
PT   A new G protein-coupled receptor gene, flh2882, useful for identifying
PT   homologues and allelic variants.
XX
PS   Claim 1; Fig 1; 80pp; English.
XX
CC   The sequence is the G protein coupled receptor flh2882 gene. The flh2882
CC   protein AAY28488 is expressed predominantly in the brain and is 337 amino
CC   acids in length. The protein has seven transmembrane domains, AAY28489-
CC   Y28495. The type of response mediated by the flh2882 protein is dependent
CC   on the type of cell in which it is expressed eg. binding of ligand to
CC   flh2882 may result in adhesion migration or differentiation. Regardless
CC   of the cellular activity, it is universal that flh2882 interacts with a
CC   "G protein". The flh2882 protein and gene are useful for isolating
CC   compounds that modulate expression of or bind and interact with flh2882.
CC   They can also be used to identify homologues and allelic variants
XX
SQ   Sequence 2581 BP; 499 A; 740 C; 753 G; 589 T; 0 U; 0 Other;
```

```
Query Match             99.5%;  Score 1009.2;  DB 2;  Length 2581;
Best Local Similarity   99.7%;  Pred. No. 8.7e-172;
Matches 1011;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

RESULT 7
ADF70549
ID   ADF70549 standard; DNA; 1728 BP.

```
Qy        1 ATGAACTCGTGGGAGGACGCGGGCCTGGGCGGGGGGCGGGCGGGCTACTGGTGGGCTCGGGCACGGATGGGGCGTCTCGGCTG 60
Db      184 ATGAACTCGTGGGAGGACGCGGGC                                                              243

Qy       61 CTGTCCAACCGCGTCGGTGGCTGG                                                              120
Db      244 CTGTCCAACCGCGGCTCGGTGTGCTTCTCCGCTCTGCTGCTGCAGCGACATCCGCGCGCCAGGCG                     303

Qy      121 CCCGGCGGCTCTTCACCCTGAACCTCACGTGCGCGGGAACCTCACAGCGCGGTGGTGGGCTGGGTCAACATG             180
Db      304 CCCGGCGGCTCTTCACCCTGAACCTCACGTGCGCGGGAA                                              363

Qy      181 CGCGGCTCACGCTGGCCGGCGTCGTGGGGGCAGCGGAGCCGACCGGCCTGTGCCGC                             240
Db      364 CGCGGCTCACGCTGGCCGGCGTCGTGGGGGCAGCGG                                                 423

Qy      241 CTGCGCTGCCTTCCTCGACACCTTCCTGGCTGCCAACTCAGCACGTGGGCCCGCGCTC                            300
Db      424 CTGCGCTGCCTTCCTCGACACCTTCCTGGCTGCC                                                   483

Qy      301 AGCAATCGACCGCTGGGGTGGGTGGGCGTGGGCGTGAGCTACCGGGGCCAAGATGCGCCTC                         360
Db      484 AGCAATCGACCGCTGGGGTG                                                                 543

Qy      361 CGGCGACGCGCGGCTCATGGTGGCGTTACACGTGGCTCTCACCGGCTCACCTTCCCAGCCGCC                       420
Db      544 CGGCGACGCGCGGCTCATGGTGGC                                                             603

Qy      421 GCGCTCGCCCTTGTCTTGGCGCTTCCACCAGTGTACCGCTCGTCGTGCACGCTGTGCAGC                          480
Db      604 GCGCTCGCCCTGTCTTCCACCAGTG                                                            663

Qy      481 CGGGCGGCGCAGCAGCGAGACGGCCTCCGCGTCCTTCGCCGTCTTCACTGGCCTGGGCTCTCAGC                     540
Db      664 CGGGCGGCGCAGCAGCGAGACGGCCTC                                                          723

Qy      541 TTCCTGCTCTCCTTCGTCGTTGCTGCAGTGGCTCAAGGTGGCCCGC                                        600
Db      724 TTCCTGCTCTCCTTCGTCGTTGCTG                                                            783

Qy      601 TTCCATTGCAAGCCCATCACGTCAACGTGACACCGCTGGTCGTGCTGGGTGACCTG                              660
Db      784 TTCCATTGCAAGCCCATCACGTCAA                                                            843

Qy      661 CACCCCAGTGTGCGGGAACCGCTGTCTGGAGGAGCGGACAGCGACC                                        720
Db      844 CACCCCAGTGTGCGGGAACCGCTGTC                                                           903

Qy      721 AAGAGATCAGCACCTTCATAGGGACCTGGGAGAGCGGACC                                              780
Db      904 AAGAGATCAGCACCTTCATAGGGAC                                                            963

Qy      781 AGGCTAGTGGAGCTCTTCTCCACCGTGCCCATCGGCTCCCA                                             840
Db      964 AGGCTAGTGGAGCTCTTCTCCACCGTG                                                         1023

Qy      841 TGCTTGGGCGTTACGCAAGGGCCATCCGACCCCTTTGTACTCCTTACTGCGACCAG                            900
Db     1024 TGCTTGGGCGTTACGCAAGGGCCATCCGACC                                                    1083

Qy      901 TACCGCAAAAGCTGCAAGGAGATTCTGAACAGAGGCTCTCGACAGACGCTCCATCCACTCC                        960
Db     1084 TACCGCAAAAGCTGCAAGGAGATTCTGAACAGAGG                                                1143

Qy      961 TCTGGCCCTCACAGCGACTCTCACAGCCAGAGACAACATTCTCGCGGTGTCTGAGTGA                          1014
Db     1144 TCTGGCCCTCACAGCGACTCTCACAGCCA                                                      1197
```

XX
AC   ADF70549;
XX
DT   12-FEB-2004 (first entry)
XX
DE   Orphan receptor ligand-related human protein gene SeqID172.
XX
KW   ligand; orphan receptor protein; fusion protein; fluorescent protein;
KW   cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
KW   GFPuv; Enhanced GFP; EGFP; human; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200307I272-A1.
XX
PD   28-AUG-2003.
XX
PF   21-FEB-2003; 2003WO-JP001901.
XX
PR   22-FEB-2002; 2002JP-00045728.
PR   23-JUL-2002; 2002JP-00213949.
PR   11-OCT-2002; 2002JP-00298237.
XX
PA   (TAKE) TAKEDA CHEM IND LTD.
XX
PI   Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
XX
DR   WPI; 2003-697654/66.
DR   N-PSDB; ADF70447.
XX
PT   Transformation of cells with a fusion protein of an orphan receptor
PT   protein with a fluorescent protein useful for identification of ligands
PT   to the orphan receptor.
XX
PS   Example 4; SEQ ID NO 172; 594pp; Japanese.
XX
CC   This invention relates to a novel method of identifying ligands to an
CC   orphan receptor protein which comprises transforming cells with DNA
CC   encoding a fusion protein of the orphan receptor with a fluorescent
CC   protein, so that the fusion protein is expressed in the cells (or cell
CC   membranes isolated from them) and contacting the cells with the potential
CC   ligand to be tested. A suitable fluorescent protein for incorporation in
CC   the fusion protein is green fluorescent protein (GFP), for example GFP-1,
CC   wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
CC   identification of ligands binding to an orphan receptor protein.
XX
SQ   Sequence 1728 BP; 382 A; 521 C; 452 G; 373 T; 0 U; 0 Other;

Query Match   99.2%; Score 1006.2; DB 10; Length 1728;
Best Local Similarity   99.7%;
Matches 1008; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy        1 ATGAACTCGTGGGAGGACGCGGGCCTGGGCGGGGGGCGGGCGGGCTACTGGTGGGCGTCTCGGCTG 60
            ||||||||||||||||||||||||
Db        1 ATGAACTCGTGGGAGGACGCGGGC                                          60

Qy       61 CTGTCCAACCGCGTCGGTGGCTGG                                          120
            |||||||||||
Db       61 CTGTCCAACCGCGGCTCGGTGTGCTTCTCCGCTCTGCTGCTGCAGCGACATCCGCGCGCCAG    120

Qy      121 CCCGGCTCACGCTGGCCGGCGTCGTGGGGGCAGCGGAGCCGACCGGCCTGTGCCGC         180
            ||||||
Db      121 CCCGGCGGCTCTTCACCCTGAACCTCACGTGCGCGGGAACCTCACAGCGCGGTGGTGGGCTGG    180

Qy      181 CCGCTCACGCTGCTGGCCGGCGTCGTGGTGCGGCAGCGGCCGACCTTGCTGTGCACCGGTC    240
            |||||||||
Db      181 CCGCTCACGCTGCTGGCCGGCGTCGTGGTGCGGCAGCGGCCGACCTTGCTGTGCACCGGTC    240

Qy      241 CTGGCTGCCTTCCTCGACACCTTCCTGGCTGCCAACTCAGCTCAGCGCATGGCTGGCCTGTC    300
            ||||||||||||||||||||
Db      241 CTGGCTGCCTTCCTCGACACCTTCCTGGCTGCCAACTCAGC                         300

Qy      301 AGCATCGACCGCTGGGTGGGTGGGCGTGGGCGTGAGCTACCGGGGCCAAGATGCGCCTC      360
Db      301 AGCATCGACCGCTGGGTGGGTGGGCGTGGGCGTGAGCTACCGGGGCCAAGATGCGCCTC      360
```

RESULT 8

ABS73406
ID   ABS73406 standard; cDNA; 1014 BP.
XX
AC   ABS73406;
XX
DT   04-DEC-2002 (first entry)
XX
DE   cDNA encoding human GPCR GPR26.
XX
KW   prostatic hypertrophy; gene; 88.
KW   prostatic hyperplasia; epilepsy; prostate cancer; rhinitis; angina;
KW   myocardial infarction; MI; stroke; glaucoma; anxiety;
KW   psychotic disorder; asthma; bronchospasm; anaesthesia;
KW   hypertension; reflux disease; depression; migraine; schizophrenia; ulcer;
DE   Human; transmembrane receptor; G-protein coupled receptor; GPCR; allergy;
OS   Homo sapiens.
XX
KW   prostatic hyperplasia; epilepsy; prostate cancer; rhinitis; angina;
KW   myocardial infarction; MI; stroke; glaucoma; anxiety;
KW   psychotic disorder; asthma; bronchospasm; anaesthesia;
KW   hypertension; reflux disease; depression; migraine; schizophrenia; ulcer;
XX
PN   WO200268600-A2.
XX
PD   06-SEP-2002.
XX
PF   26-FEB-2002; 2002WO-US005625.
XX
PR   26-FEB-2001; 2001US-0271913P.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Liaw CW, Chalmers DT, Behan DP, Maciejewski-Lenior D, Leonard JN;
PI   Lin I, Ortuno D;
XX
DR   WPI; 2002-706980/76.
DR   P-PSDB; ABG95179.
XX
PT   New human G-protein coupled receptor (GPCR), useful for screening agonist
PT   or inverse agonist compounds for treating diseases associated with GPCR.
XX
PS   Example 1; Page 198-199; 201pp; English.
XX
CC   The present invention relates to transmembrane receptors, particularly
CC   endogenous human G-protein coupled receptors (GPCRs), mutant (non-
CC   endogenous) versions of the GPCRs, and the polynucleotide sequences
CC   encoding them. The GPCRs are useful for screening agonist or inverse
CC   agonist compounds for treating diseases associated with GPCR. Diseases
CC   that can be treated with such compounds include allergies, hypertension,
CC   reflux disease, depression, migraine, schizophrenia, ulcers, psychotic
CC   disorders, asthma, bronchospasm, anaesthesia, myocardial infarction (MI),
CC   stroke, glaucoma, prostatic hypertrophy, epilepsy, prostate cancer,
CC   anxiety, prostatic hypertrophy, rhinitis, and angina. The present
CC   sequence encodes an endogenous human GPCR
XX
SQ   Sequence 1014 BP; 153 A; 364 C; 298 G; 199 T; 0 U; 0 Other;

Query Match   97.3%; Score 986.8; DB 6; Length 1014;
Best Local Similarity   98.3%; Pred. No. 8.7e-168;
Matches 997; Conservative   0; Mismatches   17; Indels   0; Gaps   0;

Qy   1 ATGAACTCGTGGGACGCCGGGCCCTGGGCGGGGGGGCTACTGGTGGGGCACGATGGGCGTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAACTCGTGGGACGCCGGGCCCTGGGCGGGGGGGCTACTGGTGGGGCACGATGGGCGTG 60

Qy   61 CTGTCCAACGGCGTGGTTGCTCTCTGCCGTGCGGGGGACAGTCCGGCGCGGCGTCCGCTG 120
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db   61 CTGTCCAACGGCGTGGTTGCTCTCTTGCCGTGCAGGCGGGGACAGTCCGGCGCGGCGTCCGCTG 120

Qy   121 CCGGGGCGCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGCTGTGTGCACGTGGTCAACATG 180
              |||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCGGGGCGCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGCTGTGTGCACGTGGTCAACATG 180

Qy   181 CCGCTCACGCTGGGCGCGTCGTGGGCGCAGGGCGCAGGGGCCGACCGCCGTGTGCCGC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGCTCACGCTGGGCGCGTCGTGGGCGCAGGGCGCAGGGGCCGACCGCCGTGTGCCGC 240

Qy   241 CTGGCCTGCCTTCCTCGACACCTTCCTGGGCTGCCAACTCCATGCTCAGCATGCCGCCTC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGGGCTGCCTTCCTCGACACCTTCCTGGGCTGCCAACTCCATGCTCAGCATGGCCGCTC 300

Qy   301 AGCATCGACACCGCTGGGGTGGCCGTGGCCGTGTGCGTGCTTCCCGGTAGCACAGATGGCGCTC 360
              ||||||||||||||||||||||||||||
Db   301 AGCATCGACACCGCTGGGGTGGCCGTGGCCGTGTGCGTGCTTCCCGGTAGCACAGATGGCCGCTC 360

Qy   361 CCGAGATGGCGCGCGCTCATCCGGCGCTACACGTGGGCGCACCTTCCCAGCCGGC 420
              |||||||||||||||||||||||||||||||||||
Db   361 CCGAGATGGCGCGCGCTCATCCGGCGCTACACGTGGGCGCACCTTCCCAGCCGGC 420

Qy   421 GCGCTCGCCCTGTCCTGGCTCGGCGTTCCACCAGTCGTACGGCGCACCGCTGTGCGAGC 480
              |||||||||||||||||||||||||||||
Db   421 GCGCTCGCCCTGTCCTGGCTCGGCGTTCCACCAGTCGTACGGCGCACCGCTGTGCGAGC 480

Qy   481 CGGCGGCCGGCGGGCGCACCGCTTGCGCCGTTATTCACTGGCGGCTTCTCAGC 540
              |||||||||||||||||||||||||||||||||
Db   481 CGGCGGCCGGCGGGCGCACCGCTTGCGCCGTTATTCACTGGCGGCTTCTCAGC 540

Qy   541 TTCCTGCTCTCCTTCGTCGTGGTGCTGCTGCACGTACCTCAAGGTGCTCAAGGTGGCCCGC 600
              |||||||||||||||||||||||||||||||||
Db   541 TTCCTGCTCTCCTTCGTCGTGGTGCTGCTGCACGTACCTCAAGGTGCTCAAGGTGGCCCGC 600

```
Qy    601  TTCATTGCAAGGCGATCAGTGATCAGCATGTCAGAGACGCTGGTGCTTGCTTGCTTGGTTGGACCTG   660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TTCATTGCAAGGCGATCAGTGATCAGCATGTCAGAGACGCTGGTGCTTGCTTGCTTGGTTGGACCTG   660

Qy    661  CACCCCAGTGTGCGGGAACCGCTGTCTGGAGGAGCAGAGACGGAAGCGGACAAGCGAGCCACC       720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  CACCCCAGTGTGCGGGAACCGCTGTCTGGAGGAGCAGAGACGGAAGCGGACAAGCGAGCCACC       720

Qy    721  AAGAAGATCAGCACCTTCATTAGGGACGACTTCCTTGTGTGTGATCAGCATGTGATCACC          780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  AAGAAGATCAGCACCTTCATTAGGGACGACTTCCTTGTGTGTGCTTCGGCGCCCTAPGTGTGATCACC  780

Qy    781  AGGCTAGTGAGCTCTTCTCCACGGTGCTCCACCTGGGCGGGTTGCTGTTCCAAG                840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  AGGCTAGTGAGCTCTTCTCCACGGTGCTCCACTGGGCGGGGTTGCTGTTCCAAG                840

Qy    841  TGCTTGGGCGTACAGCAAGGACGGCATCCAGCACCCTTTGTATACTCCTTACTGCGACACCAG       900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TGCTTGGGCGTACAGCAAGGACGGCATCCAGCACCCTTTGTGTACTCCTTACTGCGACACCAG       900

Qy    901  TACCCCAAAAGCTGCAAGAGGATTCTGCACAGACGGCTCCATCCACTCC                     960
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TACCCCAAAAGCTGCAAGAGGATTCTGCACAGACGGCTCCATCCACTCC                     960

Qy    961  TCTGGCCTCACAGGCGGACTTCACAGCCAGGAACAATTCTGCCGGGTGTCTGAGTGA            1014
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961  TCTGGCCTCACAGGCGGACTTCACAGCCAGGAACAATTCTGCCGGGTGTCTGAGTGA            1014
```

RESULT 9

AD029057

Novel mammalian G protein coupled receptors, useful for identifying
compounds that modulates diagnosing and treating disease condition
associated with GPCR dysfunction e.g. autoimmune diseases, angina
pectoris, Parkinson's disease.

XX

Claim 13; SEQ ID NO 156; 542pp; English.

XX

```
CC   The invention relates to human and mouse G protein-coupled receptors
CC   (GPCRs) and nucleic acids encoding them. The invention also relates to
CC   sequences at least 90% identical to them; methods of diagnosing diseases
CC   of the invention; methods of treating, preventing or diagnosing diseases
CC   associated with GPCRs of the invention; methods of screening for
CC   compounds useful in the treatment of GPCR-related diseases; a transgenic
CC   mouse comprising a GPCR gene of the invention; a mouse comprising a
CC   mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC   from the transgenic mice; kits comprising several mice, each of which has
CC   a mutation in a different GPCR gene of the invention; and kits comprising
CC   probes which hybridise to GPCR polynucleotides of the invention. The
CC   invention further discloses variants of the GPCR polypeptides and vectors
CC   comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC   be used in the diagnosis, treatment or prevention of a wide variety of
CC   diseases including neurological disorders (e.g., Alzheimer's disease,
CC   depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC   disorders of the adrenal gland; disorders of the colon or intestine
CC   (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC   syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC   myocardial infarction); muscular disorders; blood disorders (e.g.,
CC   anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC   AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC   arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC   obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC   diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC   uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC   thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC   nucleic acid of the invention. Note: The full sequence data for this
CC   patent did not form part of the printed specification; those sequences
CC   not shown were obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1011 BP; 165 A; 347 C; 282 G; 217 T; 0 U; 0 Other;
```

```
Query Match           85.7%;  Score 869.2;  DB 12;  Length 1011;
Best Local Similarity 91.3%;
Matches  922;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;
                     Pred. No. 1e-146;
```

```
Qy      1  ATGAACTCGTGGGACGCGGGGCGCCTGGCGGGGGCGTACTGGTGGCACGATGGGGGTCTCGCTG   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  ATGAACTCGTGGGACGCGGGGCGCCTGGCGGGGGCGTGCTGGTGGCACTATCGGCGGGGCGCTG   60

Qy     61  CTTGTCCAACGCGCGGTTGCTGTGCGCTCGCCTGCTGCAGAGCGGCGGCGACAATCCGCGCCCCCGAGGG  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  CTGTCCAACGGCGCTGGGTGCTGCTCCCTCCGGTCGCCTGACATCCGCCAGCGTGGCAGGCG   120

Qy    121  CCGGGGCGCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGCGTGTGACCGTGGTCAACATG   180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  CCGGGGCGCTCTTCACTCTCAACCTCACGTGGCCAACCTGCGTTGTGTACCGGTGGTCAACATG   180

Qy    181  CCACTCAAACACTGGCCGCGGCCGGGCGTGCGTTGCGCAGCGCCGGCGGGACCGGCCTTGTGCCGC  240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CCACTCAAACACTGGCCGCGGCCGGGCGTGCGTTGCGCAGCGCCGGCGGGACCGCCCTGTCCCGC  240

Qy    241  CTGGCGGCTGCCTTCCTCGACACCCTTACCTGGCTGGCCAACTCCATGCTCAGCTATGGGCTC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  CTGGCCGCGCCTTCCTCTGGACACCTTTCTGGCGCCAACTCCATGCTCAGCTATGGGCTC   300

Qy    301  AGCATGGACGCGCTGGGTGGGCCGTGGCCGTGGTCTTCCCGTGAGCGTTACCGGACGGCCCTC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301  AGCATCGACCCGCTGGGTGGGTCGTGGTCGTCTTCCCGCTGAGCTACCGTGCCAAGATGCGCCTC  360

Qy    361  CGCGACGCGGCGCGCTCACTGGTGGCCGGTCTTCCCGCTGAGCTTGCCACGTTCCCAGCGCGC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  CGAGATGCGCCGCGCCCTTCATTGGTGGTGCGCTTACACGTGGCTGCACGTGGCGCTCAACTTCCCAGCGCGCCACC  420
```

```
Qy    421  GCGCTCGCCCTGTCCTGGCTCGGCTCGGCTTCCACCAGCTTGGTGCGACGCTGTGCAGC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421  GCGCTTCGCCCTGTCCTGGCTCGGCTTCCACCAGCTTATGCCTTCGTGCACACTGTGCAGC  480

Qy    481  CGGCGGGCCAGAACGAGAGCGGCGGCTGCGCCTTCGCCGGTTCCACCAGCTTCTCAGC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  CGGCGGGCCGGACAGAGCGGCTGCGCTTTGCCGTGTCTTCACGCAGCCCATGGCGCTCAGC  540

Qy    541  TTCCTGCTCTTCCCTTCGTGCTGTGTTCACGTGCACGTGGTGGCCCCGC  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541  TTCCTGCTCTCCTTCATCGTGCTGTGCTTCACGTGTACCTCAAGGTTGGCCCCGC  600

Qy    601  TTCCATTGCAAGCCGCCATCGACGTTGATCACCATTGACAGGACCTG  660
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    601  TTCCACTGCAAGCCGCCATCGACGTGATCACCATTGACAGGACCGTTGCTGCTGACATA  660

Qy    661  CACCCCAGTGTGCGGGACGTGTTCTGGAGGAGCGAAGACGGAGCCACC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    661  CACCCCAGTGTGAGGGAACGGTGTCTGGAGGAACGGGAGCCAGCGTGCCACC  720

Qy    721  AAGAAGATCAGCACCTTCATAGGGACCTTTCCTTGTTGTGTTGGTTGCTCGTGGGACCTG  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721  AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTCTTTGCACCCTATGTGAGTTATACC  780

Qy    781  AGGCTAGTTGGAGCTTTCTTCTCCACGGTGCCCATCGGGTGCGCCTCCAAG  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781  AGGCGGGTGGGAACTTCTTCTCACAGCACCCATTGGCCTCTCGGGGAGTGCTGTCCAAG  840

Qy    841  TGCTTGGGGTACACGACAAGCGGCGCATCGACACCCTTGTGTACTCTTACTGGCACCAG  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841  TGCTTGGGCTTCACAGCAAGGCCGGCTTCTCGACCCCTTCGTGTATTCCTTGCTGGCACCAA  900

Qy    901  TACCGCAAAAGCTGGAAGGAGGATTCTGAACAGGCTCCTGACAGACGCTCCATCCACTCC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901  TACCGCGAGAGCTGCAAGGAGGTCCTGAACACGGAGATCTTCAAACAGACAGTCTTCCTT  960

Qy    961  TCTGGCCTCACAGGGGACTCTCACAGCCAGAACACATTCTGCCGGTGTCTGA  1010
           |||||||||||||||||||||||||||||||||||||||||||||
Db    961  GTGGGCCCTCACAGTGGACTCTCACAGGCCAGAACAATTCTGCCAGTGTGTCGGA  1010
```

RESULT 10
ID   ADQ94491   standard; cDNA; 1172 BP.
AC   ADQ94491;
XX
DT   23-SEP-2004 (first entry)
XX
DE   Rat G-protein coupled receptor sequence 115 cDNA.
XX
KW   anorectic; immunomodulator; gene therapy; body weight;
KW   mammalian sequence 115; obesity; cachexia; G protein coupled receptor;
KW   sequence 115; rat; gene; ss.
XX
OS   Rattus norvegicus.
XX
FH   Key   Location/Qualifiers
FT   CDS   27..1040
FT   /tag= a
FT   /product= "Rat G-protein coupled receptor sequence 115"
XX
PN   US200412123952-A1.
XX
PD   24-JUN-2004.
XX
PF   15-DEC-2003; 2003US-00735991.
XX
PR   23-DEC-2002; 2002US-0436375P.
XX
PA   (GOOD/) GOODNOW R A.
PA   (MARK/) MARK D F.
PA   (MART/) MARTIN M L.
PA   (ROSI/) ROSINSKI J A.
XX
PI   Goodnow RA, Mark DF, Martin ML, Rosinski JA;
XX
DR   WPI; 2004-468185/44.
DR   P-PSDB; ADQ94492.
XX
PT   Identifying compounds, useful for modulating body weight and for treating
PT   obesity and cachexia, comprises contacting a test compound with a
PT   mammalian sequence 115.
XX
PS   Example 3; SEQ ID NO 3; 28pp; English.
XX
CC   The invention describes a method of identifying compounds useful for
CC   modulating body weight, comprising contacting a test compound with a
CC   mammalian sequence 115. The method comprises: contacting a test compound
CC   with a mammalian sequence 115; determining whether the test compound
CC   binds to the mammalian sequence 115; and identifying a compound that
CC   binds to the mammalian sequence 115 as a compound useful for modulating
CC   body weight. Also described are: a pharmaceutical formulation or
CC   composition, for modulating body weight, comprising a compound that
CC   modulates the activity of a mammalian sequence 115, mixed with a
CC   pharmaceutical carrier; a package comprising the pharmaceutical
CC   formulation and instructions for administering the pharmaceutical
CC   formulation for modulating body weight; preparing a pharmaceutical
CC   composition useful for modulating body weight; treating obesity and
CC   cachexia; and an antibody that recognises an isolated polypeptide
CC   comprising a sequence of 1014 or 337 amino acids (SEQ ID NOS: 5 or 6).
CC   The method is useful for identifying compounds useful for modulating body
CC   weight. The formulation or composition and methods are useful for
CC   treating obesity and cachexia. This sequence encodes rat G protein
CC   coupled receptor sequence 115.
XX
SQ   Sequence 1172 BP; 201 A; 383 C; 338 G; 250 T; 0 U; 0 Other;

Query Match   85.6%;   Score 868.4;   DB 12;   Length 1172;
Best Local Similarity 91.0%;   Pred. No. 1.4e-146;
Matches 923; Conservative   0; Mismatches   91; Indels   0; Gaps   0;

```
Qy      1  ATGAACTCGTGGGACCGCGGGCCTGGCGCGGGGGCTGCTCGGTGGCCACCAATCGGGGTCTCGCTG  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27  ATGAACTCGTGGGACACCGGGCCTGGGGGGGGGCTGCTCGGTGGGCACAATCGGCGGGTGTCGCTG  86

Qy     61  CTGTCCAAACCGCGCTTGGTGCTGCTTCTGCCTGCACCAGCCGCGGACATCCCGCGCCAGGCG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87  CTGTCCAAACGGGCGTTGGTGCTCCTCTCCTGCACAGCGGGCCATCCCGCGCCAGGCG  146

Qy    121  CCGGGCGCTCTTCACCCTGAACCTCAGGTGCGGGAACCTGGTCAGCGTGGTCAACATG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147  CCGGGCGCTCTTCACTCTCAACCTCAGGTGCGGGAACCTGCTGTGTACCGTGGTCAACATG  206

Qy    181  CCGCTCACGCTGCGCCGCGGCAGCCGCGGGGGGGCAGCCGCGGCTGTGCCGC  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207  CCACTAACACTGGCCCGGCGCGTCGTGGGCACAAGGGCACGCCGGCCTGTCCGC  266

Qy    241  CTGGCTGCCTTCCTTGACACACTTCCTGGCTGCCAACTCCATGCCATGGGCCCGGCTC  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267  CTGGCCCGGCCCTTCCTCGACACCTTTCCTGGCCGCAACTCCAGTGTATGGGCCCGGCTC  326

Qy    301  AGCATGGCACCCGCTGGGGTGGGGCGTTCCCGGCTCAGTGCTGGCTAACGGCCAAGATCGGCTC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327  AGCATCGACCGCGCTGGGGTGGGCTGTGGTTCTTCCCGGCCAAGATGGCCGGCGTC  386

Qy    361  CGCGAGCGGCGGGCCTACACGTGGCTGCACGCGGCTCACCTTCCCAGCGCC  420
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387  CGAGATGCCGGGCCTTCATGGTGGCCTACACGTGGCCCACCTTTCCCGGCCACC  446

Qy    421  GCGCTCGCCCTGTCCTGGCTCGGCTTCCACCAGCTTGGTGCGACGCTGTGCAGC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447  GCGCTTCGCCCTGTCCTGGCTCGGCTTCCACCAGCTGTATGCCTTGCACGTGCAGC  506

Qy    481  CGGCGGGCCAGAACGAGAGCGGCGGCTGCGCCTTCGCCGGTTCCACCAGCTTCTCAGC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481  CGGCGGGCCGGACAGAGCGGCTGCGCTTTGCCGTGTCTTCACGCAGCCCATGGCGCTCAGC  540
```

```
Db   507 CGGGGAGCCCGACGAGCCCCTGGGCTTTGCTGTTCTTCACCAGGCGCTTAGC          566
         |||||||||||||||||||||||||||||| ||||||||||||||||||||||

Qy   541 TTCCTCGTCCTCTTCGTCGTCGTGCTGCACGTACCTTCAAGGTGCTCAAGGTGGCCCGGC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 TTCCTCGTCCTCTTCGTCGTCGTGCTGCACGTACCTTCAAGGTGCTCAAGGTGGCCCGT  626

Qy   601 TTCCATTGCAACGGCATCGACTGATCACCATGCAGACGCTGGTGCTGTGGTGGACCTG    660
         |||||||||||||||||||||||||||||||||||||||| ||||| ||||||||||||
Db   627 TTCCACTGCAAGCGCATCGACTGATCACCATGCAGACGCTCGTGTCGTGGTGGACCATC  686

Qy   661 CACCCCAGTGTGTGCGGGAAACCGCTGTCTGGAGGAGCAGAAGCGGAGGCGGCACCACC 720
         |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   687 CATCCCAGTGTGTGAGGGAACGAGTCTGGAGGAGCAGAAGCGGAGGCGGCACCGTGCACC 746

Qy   721 AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGCTTGCGGCCCTATGTGATCACC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   747 AAGAAGATCAGCACCTTCATAGGGACCTTCCTCTGTGCTTGCGGCCTATGTGATTACC   806

Qy   781 AGGCTAGTGGAGCTCTTCACGGGTGCCCACTGGGGGTGCTTGTCCAAG             840
         |||| ||||||||||||||||||||||||||||||||||||||||||||
Db   807 AGGCCAGGTGGAAACTCTTCTCCACAGCACCCATTGACTCAGCACTGTGCCAAG       866

Qy   961 TCTGGCCTCACGAGCGACTCTCCACGACGCAGAACATTCTGCCGGTGTGTCTGAGTGA    1014
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   927 TACCCGCAGGAGCTGCAAGGAGCTTCTGAACAGGATCTTCAACAGAGAGCGCTCCATTCACTCT 986

Db   987 GTGGGAGCTCACAGGTGACTCTTCTCCACAGCAGAACATTCTGCCCAGTGTGTCGGAATGA 1040
```

```
RESULT 11

ID   AA568630 standard; cDNA; 565 BP.
XX
AC   AA568630;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #4434.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
PR   23-AUG-2000; 2000US-00649167.
XX
PA   (HYSB-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG04443.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 1; SEQ ID NO 4434; 103pp; English.
```

The invention relates to isolated polynucleotide (I) and polypeptide (II)
sequences. (I) is useful as hybridisation probes, polymerase chain
reaction (PCR) primers, oligomers, and for chromosome and gene mapping
and in recombinant production of (II). The polynucleotides are also used
in diagnostics as expressed sequence tags for identifying expressed
genes. (I) is useful in gene therapy techniques to restore normal
activity of (II) or to treat disease states involving (II). (II) is
useful for generating antibodies against it, detecting or quantitating a
polypeptide in tissue, as molecular weight markers and as a food
supplement. (II) and its binding partners are useful in medical imaging
of sites expressing (II); (I) and (II) are useful for treating disorders
involving aberrant protein expression or biological activity. The
polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits to assess biodiversity
and to produce other types of data and products dependent on DNA and
amino acid sequences. AA564197-AAS94564 represent novel human diagnostic
coding sequences of the invention. Note: The sequence data for this
patent did not appear in the printed specification, but was obtained in
electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

```
XX
SQ   Sequence 565 BP; 190 A; 157 G; 118 T; 0 U; 0 Other;

Query Match        54.9%;     Score 556.6;  DB 5;   Length 565;
Best Local Similarity   99.3%;   Pred. No. 1.1e-90;
Matches 559;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

Qy   452 AGCTGTACGGCCTCGTGCACGCGTGCAGCGCGGGGCCGGGCGACGAGGCGCTTCGCG    511
         ||||||||||||||||||||| || || |  |||| |||||| || ||||||| ||||
Db     3 AGCTGTACGGCCTCGTGCACGGCGTGCGAGCCCGAGGCCGGGCCGCGCGCTTCCCG     62

Qy   512 TCTTCACTGGGCGCCCTTCCACCGCTCTCAGCTTCCTCGTCCTCTGTCGTCGTGCTGCA  571
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63 TCTTCACTGGGCGCCCTTCCACCGCTCTCAGCTTCCTCGTCCTCTGTCGTCGTGCTGCA  122

Qy   572 CGTACCTTCAAGGTGCTCAAGGTGGCCCGGCTTCCATTGCAACGGCATCGACTGATCACCA 631
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 CGTACCTTCAAGGTGCTCAAGGTGGCCCGGCTTCCATTGCAACGGCATCGACTGATCACCA 182

Qy   632 TGCAGACGCTGGTGCTGTGGTGGACCTGCACCCCAGTGTGTGGGGGACGCGTGTCTGGAGG 691
         ||||||||| ||||| ||||||||||||||||||||||||||||||||||||||||||||
Db   183 TGCAGACGCTCGTGTCGTGGTGGACCTGCACCCCAGTGTGTGGGGGACGCGTGTCTGGAGG 242

Qy   692 AGCAGAAGCGGAGGCGCCACAGGGAGCGCACCAAGAAAGATCAGCACCTTCATAGGGACCTTCC 751
         |||||||||||||||  || |  ||||||||||||| |||||||||||||||||||||||||
Db   243 AGCAGAAGCGGAGGCGGCGCCACCAGTGGAGCGCACCAAGAAGATCAGCACCTTCATAGGGACCTTCC 302

Qy   752 TTGTGTGCTTGCGGCGCCCTATGTGATCACCAGGCTAGTGGAGCTCTTCACGGGTGCCCA  811
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 TTGTGTGCTTGCGGCGCCCTATGTGATCACCAGGCTAGTGGAGCTCTTCACGGGTGCCA   362

Qy   812 TCGGCTCCCACTGGGGGGGGGGGGGTGCTTGTCCAAGTCGTACAGCAAGGCGCCGGACACC 871
         ||||||||||||||||||    |||||||||||||||||||||||||||  || || ||||
Db   363 TCGGCTCCCACTGGGGGGTGCTTGTCCAAGTCGTACAGCAAGGCCGCCATCGACC       422

Qy   872 CCTTTGTGTACTTCCTTTACTCGGACACCAGTACCGCAAAAGCTGCAAGGAGATTCTGAACA 931
         ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db   423 CCTTTGTGTACTTCCTTTACTCGGACACCAGTACCGCAAAAAGCTGCAAGGAGATTCTGAACA 482

Qy   932 GGCTCCTGCACAGAGACGCTTCCATCCACTCCTTCGGCGACTCTCACAGCGA          991
         ||||||||||||||||||| ||||||||||||||||||   |||||||||| ||
Db   483 GGCTCCTGCACAGAGACGCTCCATCCACTCCTTCGGCCTCACAGCGACTCTCACAGCCAGA  542

Qy   992 ACATTCTGCCCGGTGTCTGAGTGA                                     1014
         ||||||| |||||||||||||||||
Db   543 ACATTCTTGCCCGGTGTCTGAGTGA                                     565

RESULT 12

ID   ACH39042 standard; cDNA; 585 BP.
```

XX
AC   ACH39042;
XX
DT   13-OCT-2003 (first entry)
XX
DE   Human foetal brain cDNA #409.
XX
KW   Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW   genome mapping; biodiversity; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   US20030073623-A1.
XX
PD   17-APR-2003.
XX
PF   30-JUL-2001; 2001US-00918995.
XX
PR   30-JUL-2001; 2001US-00918995.
XX
PA   (DRMA/) DRMANAC R T.
PA   (LABA/) LABAT I.
PA   (STAC/) STACHE-CRAIN B.
PA   (DICK/) DICKSON M C.
PA   (JONE/) JONES L W.
XX
PI   Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR   WPI; 2003-615964/58.
XX
PS   Claim 1; SEQ ID NO 26254; 44pp; English.
XX
CC   The invention relates to an isolated polynucleotide comprising any one of
CC   38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC   determined by the technique of SBH (sequencing by hybridisation). Also
CC   included is a purified polypeptide comprising a sequence corresponding to
CC   a reading frame of the novel polynucleotide. The nucleic acid sequences
CC   are useful in diagnostics as expressed sequence tags (EST) for
CC   identifying expressed genes or for physical mapping of the human genome,
CC   in forensics, in assessing biodiversities, or in identifying mutations
CC   responsible for genetic disorders and other traits. The nucleotide
CC   sequences are also useful as hybridisation probes, as oligomers for PCR,
CC   for chromosome and gene mapping, in the recombinant production of
CC   protein, or in generating antisense DNA or RNA. The purified polypeptide
CC   is useful for generating antibodies specific for it. The present sequence
CC   is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC   for this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ   Sequence 585 BP; 118 A; 182 C; 161 G; 118 T; 0 U; 6 Other;

Query Match   46.6%; Score 472.2; DB 9; Length 585;
Best Local Similarity   99.4%;
Matches 474; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy   527 TCCAGCCTCTCAGCTTCCTGCTCTCGTTCGTTGCTGCTGCACGTACCTCAAGGTGC 586
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    49 TCCAGCCTCTCAGCTTCCTGCTCTCGTTCGTTGCTGCTGCACGTACCTCAAGGTGC 108

Qy   587 TCAAGGTGGCCCGCTTCCCATTGCAAGCGCATCGACGTGATCACCATGCACGCTCGGTGC 646
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   109 TCAAGGTGGCCCGCTTCCATTGCAAGCGCATCGACGTGATCACCATGCACGCTCGGTGC 168

Qy   647 TGCTGGTGGACCTGCACCCCAGTGTGGGGGAACGGCAGAGAAGCGGAGGC 706
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   169 TGCTGGTGGACCTGCACCCCAGTGTGGGGGAACGGCAGAGAAGCGGAGGC 228

Qy   707 GACAGCGGAGCCCACCAAGAAGAGATCAGCACACCTTCATAGGAGGACCTTCCTTGTTGTTGTCGCGC 766
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   229 GACAGCGGAGCCCACCAAGAAGAGATCAGCACACCTTCATAGGAGGACCTTCCTTGTTGTTGTCGCGC 288

Qy   767 CCTATGTGATCACCAGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGGTCCCACTGGG 826
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   289 CCTATGTGATCACCAGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGGTCCCACTGGG 348

Qy   827 GGGTGCTGTTGTCCAAGTGCTTGGCGTACAGCAGGCAAGGCCGCCATCCGACCCCCTTTTGTTGTACTCCT 886
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   349 GGGTGCTGTTGTCCAAGTGCTTGGCGTACAGCAGGCAAGGCCGCCATCCGACCCCCTTTTGTTGTACTCCT 408

Qy   887 TACTGCGGACACCAGGTACCGCAAAAGGTGCAAGGAGGAGATTCTGAACAGGCTCCTTGCTCAGAC 946
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   409 TACTGCGGACACCAGGTACCGCAAAAGGTGCAAGGAGGAGATTCTGAACAGGCTCCTTGCACAGAC 468

Qy   947 GCTCCATCCACTCCTTCTGGGCCTCACAGGCGGACTCTCACAGGACAGCAGGAACAGTTCTGCCGG 1003
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   469 GCTCCATCCACTCCTTCTGGGCCTCACAGGCGGACTCTCACAGGACAGCAGGAACAGTTCTGCCGG 525
```

RESULT 13

AAS07950
ID   AAS07950 standard; cDNA; 1092 BP.
XX
AC   AAS07950;
XX
DT   23-OCT-2001 (first entry)
XX
DE   Human cDNA encoding G-protein coupled receptor, hRUP23.
XX
KW   Human; G-protein coupled receptor; GPCR; hRUP23; agonist;
KW   inverse agonist; lung cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200136471-A2.
XX
PD   25-MAY-2001.
XX
PF   16-NOV-2000; 2000WO-US031509.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..1092
FT   /*tag= a
FT   /product= "hRUP23"
XX
PR   17-NOV-1999; 99US-0160808P.
PR   17-NOV-1999; 99US-0166099P.
PR   17-NOV-1999; 99US-0166369P.
PR   23-DEC-1999; 99US-0171900P.
PR   23-DEC-1999; 99US-0171901P.
PR   23-DEC-1999; 99US-0171902P.
PR   11-FEB-2000; 2000US-0181749P.
PR   14-MAR-2000; 2000US-0189258P.
PR   14-MAR-2000; 2000US-0189259P.
PR   10-APR-2000; 2000US-0195898P.
PR   10-APR-2000; 2000US-0195899P.
PR   28-APR-2000; 2000US-0196078P.
PR   12-MAY-2000; 2000US-0204419P.
PR   12-JUN-2000; 2000US-0203630P.
PR   12-JUN-2000; 2000US-0210741P.
PR   12-JUN-2000; 2000US-0210982P.
PR   21-AUG-2000; 2000US-0226760P.
PR   26-SEP-2000; 2000US-0235418P.
PR   26-SEP-2000; 2000US-0235779P.
PR   26-OCT-2000; 2000US-0237779P.
PR   20-OCT-2000; 2000US-0242332P.
PR   20-OCT-2000; 2000US-0242343P.
PR   24-OCT-2000; 2000US-0243019P.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Chen R, Dang HT, Lowitz KP;

```
XX
DR   WPI; 2001-355616/37.
DR   P-PSDB; AAU04377.
XX
PT   Endogenous and non-endogenous versions of human G-protein coupled
PT   receptors for direct identification of candidate compounds as agonists,
PT   inverse agonists or partial agonists for use as therapeutic agents.
XX
PS   Example 1; Page 117; 160pp; English.
XX
CC   The sequence encodes a human G-protein coupled receptor (GPCR), hRUP23.
CC   The endogenous and non-endogenous, constitutively activated versions of
CC   human G-protein coupled receptors (GPCR), are useful for direct
CC   identification of candidate compounds as receptor agonists, inverse
CC   agonists or partial agonists having applicability as therapeutic agents
CC   for treating diseases related to GPCR, e.g. lung cancer. Non-endogenous
CC   version of human GPCRs are also utilized in research settings and in
CC   vitro and in vivo system, incorporating GPCRs can be utilised to
CC   elucidate and understand the roles these receptors play in the human
CC   condition, both normal and diseased
XX
SQ   Sequence 1092 BP; 151 A; 397 C; 340 G; 204 T; 0 U; 0 Other;

Query Match             41.2%;  Score 418.2;  DB 5;  Length 1092;
Best Local Similarity   64.5%;
Matches  624;  Conservative    0;  Mismatches 343;  Indels    0;  Gaps    0;

Qy        1 ATGAACTGTGGGACCACCGGCGCCCTGGGGCTTACTCGTGGGGCCACGATGGGGTCTCGCTG   60
            ||||||||||
Db        1 ATGGGCCCCCGAGCCGTTCCGGCGGGGTCTTCTGGCGCGGGGTCTGGGCAC            60

Qy       61 CTGTCCAACGGCGGCTGGTGTGCTCGTCTCCCTGCTGCACAGCCGACATTCCGCCCCAGGCG  120
            ||||||||||          |||||    |||
Db       61 CTATCCAACGGCACTGGTGCTCGCTTTGTTGCGCCTACAGCGGTGAGCTTCCGCCACTCGAGCC 120

Qy      121 CCGGGGGCCTCTTCACCCTGACCTCACGTGGGGGAACCTGCTGTGGGTCAACATG         180
            |||||                 |||||  |||
Db      121 TCAGGGCGTTCCTCCTGTGAATCTGTCGGTCCGGCCACTCGGCGCCTGACATG           180

Qy      181 CCGGCTCACGTGGGCCGCGTCGTGGGGGGCCAGGGGGGGGACCGGCTCGTGCCGC         240
            |||||||||
Db      181 CCCTTCACGCTGCTCGGGTGGTGATGCGGGCGGGACACCGTGCCTCCCGGCTC            240

Qy      241 CTGGCTGCCTTCCTGCCACACCTCCTGGCTGTGCGGGGCGGGACGGCGCCCGCGGCTC      300
            |||||        ||     ||
Db      241 GTCATTGGCTTCCTGGACACACCTTCCTGGCGTCCAAGCGGGTGGGGGGCTG             300

Qy      301 AGCATCGACGCGCTGGGGGTGGTGGGCCGTGGTGTCTTCCCGCGTTGAGCGGCCAAGAGATGCGCCCTC 360
            |||||     ||
Db      301 AGCGCAGACCAGTGGCGTGGGCCTTCCCACTGCGCTGCACGCCGGACCTGCCACCG         360

Qy      361 CGCGACGCGGGGCCTCATGGTGGCGGCCTACACGTGGGTCGCACGGCGCCCAGCCGCC      420
            |||||        ||
Db      361 CGCTATGCCGGCGCCTGCTGCTGCGGGGTGGCTGGCCCTTTCTCAGGCGCGCT           420

Qy      421 GGCGCTCGCCCTGTCCTGGGCTCCACCAGCTGTACGCGCCTCGTGCAGGCTGTCAGC       480
            |||||        ||        |||
Db      421 GCACTTGGCGTCGTGGCTTGGGCTACAGAGCGGCCGCTTCCGGCGTTCTGTTGCTGCCCTG    480

Qy      481 CGGGGGCCCAGACGAGGCGCCTGCGCTTCGCGCCGCTTCTCACTCGGCGCTCTCAGC       540
            |||||
Db      481 CCGCCCGAGCCTGAGCGCTTCGCAGCCTTCACCCGCACGCTCCATGCCGTGGGGC          540

Qy      541 TTCCTGCTCCTTTCGTGTGTGCGCACGTACCTCAAGGTCTCAAGGTGCCCGC            600
            |||       ||
Db      541 TTCGTGCTCGCCCGCGTTGGCTCTGCCTCCAGGTCGCCACCGGGTCGCACGC             600

Qy      601 TTCCATTGCAAGGCATCGAGTGATCACCATCAGAGACGCTGGGTCGTTCTGGTGACCTG      660
            |||         ||
Db      601 AGCCACTGCCAAGCGCAATGGACACCGTCACCATGAAGGCGGCTCGCGCCGACCTG         660

Qy      661 CACCCCAGTGTGTCGGGGAAACCTGCTGTCTGGAGGAGGAGGCGGCGGCCCGCCGCCACC     720
            |||||
Db      661 AGGAAGATCAGGACACCTTCCTTGTGTGTGCTTCGGCGGCCCTATGTGATCAC            720
```

```
Db      661 AGGAAGATTGGCATTGCTATTGCGACCTTCCCATCTCATCGTGGTGTGCCTG            720

Qy      721 AAGAAGATCAGGACATCCATAGGGGACCTTCCTTGTGTGTGCTTCGGCGGCCCTATGTGATCACC 780
            |||||||
Db      721 AGAAAGATTGGCATTGCTATTGCGACCTTCCCATCTCATCGTGGTGTGCCTG            780

Qy      781 AGGCTAGTGGAGGCTCTTCTCCACAGGGTGCCCATCGGGTCCCACTGGGGGGGGGGGTCCAAG  840
            ||||
Db      781 AGGCTGGGGGGAGCTCGTGCCCTTGGGTCACCGTGAACGCCCAGTGGGGCCATCCTCAGCAAG   840

Qy      841 TGCTTGCGTGACAGCAAGGCGCCGCCATCGACCCCCTTCCTTTGTGTACTCTTACTGGGAACACCAG 900
            ||||
Db      841 TGCCTGACCTACAGAGCAAGGCGCCGGACCCCGTTCACGTACTCTCTGCCCGGCGG          900

Qy      901 TACCGGCAAAAGCTGCAAGGAGAGAATTCTGAAACAGGCCTCCTGCACAGAGACGCGCTTCCATCGCACCTCC 960
            ||||
Db      901 TTCCGGCCAAGTCCTGGGCGCGGCCATCGGGTTGCACCCCGGCTGTGAACGAGAACCCCGCGGCCCGCCAGCA   960

Qy      961 TCTGGCC 967
            |||  ||
Db      961 TCCACCC 967
```

```
RESULT 14
ADL96477
ID   ADL96477 standard; cDNA; 1092 BP.
XX
AC   ADL96477;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human G protein-coupled receptor (GPCR) cDNA #16.
XX
KW   Human; G protein-coupled receptor; GPCR; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003166148-A1.
XX
PD   04-SEP-2003.
XX
PF   16-DEC-2002; 2002US-00321807.
XX
PR   13-OCT-1998;  98US-00170496.
PR   16-NOV-2000; 2000US-00714008.
XX
PA   (CHEN/) CHEN R.
PA   (DANG/) DANG H T.
PA   (LOWI/) LOWITZ K P.
XX
PI   Chen R,  Dang HT,  Lowitz KP;
XX
DR   WPI; 2003-898073/82.
DR   P-PSDB; ADL96478.
XX
PT   New G protein-coupled receptor (GPCR), useful for preparing a composition
PT   for identifying compounds as receptors, inverse agonists or partial
PT   agonists having potential applicability as therapeutic agents.
XX
PS   Example 1; SEQ ID NO 31; 82pp; English.
XX
CC   The invention relates to G protein-coupled receptor (GPCR) polypeptides
CC   and the polynucleotides encoding them. The GPCR polypeptides are useful
CC   for preparing a composition for identifying compounds as receptors,
CC   inverse agonists or partial agonists, having potential applicability as
CC   therapeutic agents. This sequence represents human GPCR cDNA of the
CC   invention.
XX
SQ   Sequence 1092 BP; 151 A; 397 C; 340 G; 204 T; 0 U; 0 Other;

Query Match             41.2%;  Score 418.2;  DB 11;  Length 1092;
Best Local Similarity   64.5%;
Matches  624;  Conservative    0;  Mismatches 343;  Indels    0;  Gaps    0;
```

RESULT 15

ID   ABZ24430 standard; cDNA; 5665 BP.

```
Qy      1 ATGAACTCGTGGGACGGCGGGCCTGGCGGGGGTGGGCACCGATGGGCGTCTCGCTG 60
                                                          ||||||
Db      1 ATGGGCGCCCGGGCGAGGCGCTGCTGCGGGGGTCCTCCTGGGTGATGGGTACTGGGCTG 60

Qy     61 CTGTCCAACGCACGCGCGTTGGGTGCTGCTTGGCCGTGCACAGCGGCGGCATCCGCCAGGCG 120
           |||||||                                                  ||||
Db     61 CTATCCAACGCACTGGGTGCTGCTTTGTTGGCCTACAGCCGTCCCGCACTCGAGCC 120

Qy    121 CCGGGCGGCTCCTTCTTCACCCTGGTGAATCTGTTGTCTCCGCGGGGGCCAGCACATG 180
               ||||||||||||||||||||||||||||||
Db    121 TCAGGCGGTCCTCCTGGTGAAATCTGTTGCGCGTGCGGCGGCTGCGCTGGACATG 180

Qy    181 CTGGCTGCCTTCCTGGACACCTTCCTGGCTGCCAACTCCATGGCTCAGCATGGCCGCGCTC 240
               |||||||||||||||
Db    181 GTCATTGGCTTCCTGACACCTTCCTGGACACGTCCGGCGCCCGGGCGCCCCGGCCAA 240

Qy    241 CTGGGCTGCCTTCCTGGACACTTCCTGGCTGCCAACTCCATGGCTCAGCATGGCCGCGCTC 300
                         ||||||||
Db    241 GCGCTCGGCCCCTGTCTCCGGCCTTCCACCCAGCGTGGCCGGCTCGTGTGCAGC 300

Qy    301 AGCATGACGACCGCGTGGGGTGGGCCGGGTCGTCGGGCTGAAGATGCCCCTC 360
           |||||||||
Db    301 AGCGGCAGACCAGTGGCTGCCGAGTGGCCTTCCCACTGCGGTTCCCAGCCTGCGGACCG 360

Qy    361 CGCGGACGCGGGCGGGCTCATGTGGTCGGCGCTACAGTGGGCTGCACCGCTTCCCAGCGCC 420
           |||||||||||
Db    361 CGCTATGCCGGCGCCTGCTCGTGTGGGGACAGTCGTGCCTTCTCAGGCGCT 420

Qy    421 GCGCTTCGCCCTGTCTCGGGCTTCCACCCAGCGTTGTACGGCCTTCGTGTGTGCAGC 480
                    ||||||||||
Db    421 GCACTTGGCGCTCGTCTGGGCTTTGGGCTTACAGAGGCGCTTCGGGTTCGCTGCCCTG 480

Qy    481 CGGCGGCCCAGACGAGGCGCCTGGCGCTCGCGGCCTTCTTCAGGCTCTCAGC 540
           |||||
Db    481 CCGCGCCGGAGCCTGAGCGGTCGGCGCTTCGCCAGCGCCTCAGTGGCCGTGGGGC 540

Qy    541 TTCCTGCTCCTCTTCGTGGTGTCTGCACGTACCTCAAGGTGCTGGGCCGC 600
           |||||
Db    541 TTCGTGTCGGCCTGCGGCGGTCTCGCCCTCAGGTGCACCGGGGTGGGCAGCGC 600

Qy    601 AGCCACTGCCAGGCAATGGGACCACCGTCACCATGAAGAGCGGCTTCGGCGGGACCTG 660
               |||||||||||||||||||||
Db    601 TTCCAGTTGCAAGGCGCATGCAATCGAGCGTGATCACACGCGCTTGTGCTGGTGGACCTG 660

Qy    661 CACCCCAGTGTGCGGGAAGCTGTTCTGGAGGAGCGGCAGAAGCGGGGAGGCGGACGAGGGCCACC 720
           |||||||||
Db    661 CACCCCAGTGTGCGGACCTGCCTCAGCAGCAGAAAGCGCGGCGGCGCCCCGCGCACC 720

Qy    721 AAGAAGATCAGCAGACCTTCATAGGGAGACCTTCCTGGTGTCGGCTTCGGCGCCTATGTGATCACC 780
                 ||||||||||
Db    721 AGGAAGAATTGGCCATTGGCTATTGGCGACCTTTGCCCGGTATGTCATTGCATGACC 780

Qy    781 AGGCTAGTGTGGAGCTCTTCTCCACGCGGTGCCCACACTGGGGGTTGCCAAG 840
           |||||
Db    781 AGGCTGGCGGAGCTCGTGCCCTTCGTCACCGTGAACCGCCGTTCCTCCAGCGAAG 840

Qy    841 TGCTTGGCGTACGAGAAGGCCCCATCGGACACCCTTTGTGTACTCCTTACTGCGGACACCAG 900
           |||||
Db    841 TGCCTGACCTACAGCAAGGCCGGTGCCCTTCGTCACCCTGTCACTGGGGCATCAGGCAAG 900

Qy    901 TACCGCAAAAAGCTGCAAGGAGATTCTGCACAGACGCTCCATTCCACTCC 960
           ||||
Db    901 TTCCGGCCAAGTCCTGGGCCGCCATGGGTGCACCCCGGCTGTGGAAGAGAACCCCCAGCA 960

Qy    961 TCTGGGCC 967
           ||
Db    961 TCCACCC 967
```

Query Match   41.1%; Score 416.6; DB 10; Length 5665;
Best Local Similarity 64.4%; Pred. No. 1.4e-65;
Matches 623; Conservative 0; Mismatches 344; Indels 0; Gaps 0;

```
Qy      1 ATGAACTGTGGGACGGCGGGGCCTGGGCGGGGCTACTGGTGGGCACGATGGGGGTCTCGCTG 60
            ||||   |||  | || ||||| ||||||||||||    || | ||  ||| |   ||||||
Db   1892 ATGGGCCCCGGCGAGGGCGGCTGGGCGGGGGGTTCCTCGGTGATGGTACTGGGCCGGCTG 1951

Qy     61 CTTGTCCAACGGGCGGGTCGTGTGCTCTGCTCTGCCTCTGTGCCAGCGGGGACATCCCGCGCCAGGCG 120
            || ||   ||| | || || || ||||||||| ||||||||  | |||  ||  |||||||||||
Db   1952 CTATCCAAACGCACTGGTGCTTGCTTTGTTGCCGCCTACAGCGTCAGAGGCTGAGCTCGAGCC 2011

Qy    121 CCGGCGGCCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGGTGCACGTGGTCAACATG 180
            | |   || |||||||| | || |||| |  ||||||||  | |||| || |||  |||
Db   2012 TCAGGCGGTCTCCTCTGTGAATCTGTCTCTGGGCCACCTGCTGCTCTGGAACATG 2071

Qy    181 CCGGCTCACGCTGGCCGCCGGCTGGGGCCAGCGGGCAGCGGGGGCACCGCCCTGTGCCGC 240
            || || ||||| | |||| || |||| ||||||||||| ||||| ||||| | ||| ||
Db   2072 CCCTTCACGGCTGCTCGGGTTGCATGCGCGGGGGGACACCCGTCGCGGGGGGCCATGCCAA 2131

Qy    241 CTGGGCTGCCTTCCTCGACACCTTCCTCTGGCCCAACTCCATGCTCAGCATGGGCCCGCTC 300
            |||| ||||||   | |||| | ||||| ||||||| |||||||||||||||||||||
Db   2132 GTCATTGGCTTCCTGGACACCTTCCTGGGGTCCAACCGCGGGCTCCAAGCGGCGTGGCCCTG 2191

Qy    301 AGCAATCGACCCCTGGGGTGGCCGGTGGGTCTTCCCGCTGAGCTACCGGGGCCAAGATGCGCCTC 360
            ||| | |||| |||| || ||| |||| || | |||||  ||||||| |||| |||| ||
Db   2192 AGCGCGACCAGTGGCTGCAGTGGGCTTCCCACTGCGCTACGCGGGGGGCTGAGGGTGGCGGGCCTG 2251

Qy    361 CGCGACGGCGGGCTCATGGTGGCCTACACGGTGGCGTCGCACGGCGGCCTCACCTTCCCAGCGCCGCC 420
            ||| | |||| || ||  | ||| ||| ||||| | || |||  ||| ||||||||||||||
Db   2252 CGCTATGCCGGCCTGCTGCGGGGCGTCGGGGGGACACCGTCGTGGCCTTCTCAGGCGCT 2311

Qy    421 GCGCTTCGCCCCTGTCTCTGGGCTTCCACCAGCTGTAGGCGCCTCGTTGCCACGCTGTGCAGC 480
            || || |||| | |||| |||| |||| ||||| ||| ||| | | |  |  |||| ||||||
Db   2312 GCACTTGGGCTGCTCGTGGCTTGGCCTTACGACAGGCGCTTCCTCGTTCGTTGCTGCCCTG 2371

Qy    481 CGGCGGCCCAGACGAGGGCCTGGGCGTTCGCCCGTCTTCACTGGGCGCCTTCCAGCTCTTCAGC 540
            |||| || |||| ||| | |||| | | || ||| | |||| ||| | ||||||||||| |||
Db   2372 CCGGCCGGAGCCTGAGGGGTTCCGCGGCTTCAGCCTTCACCCGCCATGCCGTTGGGGC 2431

Qy    541 TTCTGCTCTCCTTCGTCGTCTGCTCTCGCACGTACCTCAAAGGTGGTTGGCCCCGC 600
            ||| | |||| ||| | ||| |  |||||  |||| |||||| |||| ||||||
Db   2432 TTCGTTGCCTGCCGCTGGTCTCTGCCTCACCCGCAGGGTGCACCCGGGTTGGCAGC 2491

Qy    601 TTCCATTGCAAGCGCCATCGACGTGATCACCATGCAGACGCTGGTGGTGCTGGTGGGACCTG 660
            |||||| | ||| ||||| |   |||||||||||||||||| | || ||| || ||||||
Db   2492 AGCCACTGCCAGCGCATGGACACCCGTCACCATGAAGGGCTCGCCGCTCGCGCCGCAGCCTG 2551

Qy    661 CACCCCAGTGTGCGGGGAACGGCTGTCTGGAGGAGCAGAAGCGGACGACGCACGAGCCACC 720
            |||||| |||| |  |||| | |||| ||||| | ||||| |  || |||||| |||||||
Db   2552 CACCCCAGTGTGCGGCAGCGCTGCCTCATCCAGCGAGGAGGCGCGCGCGTCTCGCGACCTG 2611

Qy    721 AAGAAGATCAGCACCTTCAATAGGGGACCTTCCTTGTGTTGCTTCGCGCCCTATGTGATCACC 780
            || | |||||  ||| |   | || || ||||||||||| |||  | ||||  ||| ||||
Db   2612 AGGAAGATTGGCAATTGCTATTGCGACACCTTCCTCATCTGCTTTGCCCCGTATGTCATTGACC 2671

Qy    781 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATGGCTCCCAACTGGGGGTTGCTGTCCAAG 840
            |||| || ||||| |||||||| ||| | ||| ||||||||||| |||| || | ||||
Db   2672 AGGCTGGCGGAGCTCGTGCCCTTCGTCACCGGACGCCCAGTGGGGGCAACCCTCAGCAAG 2731

Qy    841 TGCTTGGGGTACAGCAAAGGCCGCATCGACCCCTTTGTGTACTCCTTACTGCGACACCAG 900
            ||| ||||| ||| ||  ||||||| |||| |  |||| |||||||||||||||||||
Db   2732 TGCCTGACCTACACAGAGAAGGGGGTTGCCGACCGTTCACGTACTCTTCTGCTCCGCGGCCG 2791

Qy    901 TACCCGCAAAAGCTGCAAGGAGATTCTGACACAGGCTCCATCCACTCC 960
            ||| ||| || ||| |||| ||| | |  ||| |||| | ||||||
Db   2792 TTCCGCCAAGTCCTGGCCGGCATGGTGCACCGGCTGCTGAAGAGAACCCGCGCCCAGCA 2851

Qy    961 TCTGGCC 967
            ||| |||
Db   2852 TCCACCC 2858
```

THIS PAGE BLANK (USPTO)

Bowman, A.
10/735991
Seg. ID 5

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: July 5, 2005, 09:08:04 ; Search time 4630 Seconds
               (without alignments)
               10612.006 Million cell updates/sec

Title: US-10-735-991-5
Perfect score: 1014
Sequence: 1 atgaactcgtgggacgcgggg..........ttctgcggtgtctgagtga 1014
Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0
Searched: 4708233 seqs, 2422607955 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :

        GenEmbl:*
        1: gb_ba:*
        2: gb_htg:*
        3: gb_in:*
        4: gb_om:*
        5: gb_ov:*
        6: gb_pat:*
        7: gb_ph:*
        8: gb_pl:*
        9: gb_pr:*
       10: gb_ro:*
       11: gb_sts:*
       12: gb_sy:*
       13: gb_un:*
       14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	Match	DB ID	Description
1	1014	100.0	1014	6		CQ831055	CQ831055 Sequence
2	1011	99.7	1011	6		CQ831051	CQ831051 Sequence
3	1010.8	99.7	7512	6		HSA505757	Homo sapi
4	1009.2	99.5	2581	6		AJ505757	Homo sapi
5	1006.2	99.2	1011	6		BD130531	Ligand re
6	967.4	95.4	969	6		BD130532	Ligand re
7	868.4	85.6	1172	6		CQ729494	Sequence
8	868.4	85.6	1172	6		CQ831053	Sequence
9	868.6	65.9	196808	10		AF208288	Rattus no
10	668.6	65.9	196808	9		AC009987	Homo sapi
11	589.6	58.1	162502	10		AC105062	Mus muscu
12	579	57.1	68539	2		AC099826	Papio ham
13	456.2	45.0	525	10		AY255585	Mus muscu
14	418.2	41.2	1092	6		AX148190	Sequence
15	416.6	41.1	1092	6		AX592452	Sequence
16	416.6	41.1	1352	6		AX592450	Sequence
17	415	40.9	1029	6		CQ716319	Sequence
18	415	40.9	1092	6		AX085586	Sequence
19	415	40.9	1092	6		AX148254	Sequence
20	415	40.9	1092	6		AX168152	Sequence
20	415	40.9	1092	6		AX316190	Sequence
21	415	40.9	1092	6		AX670876	Sequence
22	415	40.9	1092	6		BD095690	Novel gua
23	415	40.9	1092	6		BC069343	Homo sapi
24	415	40.9	1092	6		BC069813	Homo sapi
25	415	40.9	1188	6		AX147940	Sequence
26	415	40.9	1188	6		AX521989	Sequence
27	415	40.9	1510	6		AX376522	Sequence
28	415	40.9	1510	6		AY359107	Homo sapi
29	415	40.9	1953	6		BC057778	Homo sapi
30	415	40.9	3119	6		BC057778	Sequence
31	410.2	40.5	1092	6		CQ850655	Sequence
32	410.2	40.5	1092	6		AK128807	Homo sapi
33	410.2	40.5	1092	6		AX549269	Homo sapi
34	396	39.1	1361	9		AF411107	Homo sapi
35	298.2	29.4	181601	2		BX670878	Homo sapi
35	282.4	27.9	1000	9		BX927351	Danio rer
36	282.4	27.9	1059	9		BX670870	Sequence
37	282.4	27.9	1059	9		AX375436	Sequence
38	282.4	27.9	6784	9		AX646681	Sequence
39	282.4	27.9	6784	9		AB065673	Homo sapi
40	282.4	27.9	7020	9		AX316189	Sequence
41	282.4	27.9	196902	2		AC105345	Homo sapi
42	282.4	27.9	200000	2		AC007104	Homo sapi
43	282.4	27.9	214284	9		AC116643	Homo sapi
44	282.2	27.8	969	6		AX657446	Sequence
45	267.2	26.4	657	6		AX147758	Sequence

ALIGNMENTS

RESULT 1
CQ831055

LOCUS   CQ831055   1014 bp   DNA   linear   PAT 29-JUL-2004
DEFINITION Sequence 5 from Patent EP1437595.
ACCESSION   CQ831055
VERSION   CQ831055.1 GI:50831180
KEYWORDS
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Godnow,J.R., Mark,D.F., Martin,M.L. and Rosinski,J.A.
  TITLE   Orphan GPCR gene 115 for obesity indication
  JOURNAL   Patent: EP 1437595-A 5 14-JUL-2004;
            F. HOFFMANN-LA ROCHE AG (CH)
FEATURES   Location/Qualifiers
     source   1..1014
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"

ORIGIN

Query Match   100.0%; Score 1014; DB 6; Length 1014;
Best Local Similarity   100.0%;
Matches 1014; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 ATGAACTCGTGGGACGCGGGGCCGCGGCGCCTGGCGGGGGGCTACTGGTGGGCGGTCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAACTCGTGGGACGCGGGGCCGCGGCGCCTGGCGGGGGGCTACTGGTGGGCGGTCGCTG 60

Qy   61 CTGTCCAACGGCGCTGGTTGCTCCTCTGCCTGCTGCACGCGGGACAATCCGCGCCAGGCG 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGTCCAACGGCGCTGGTTGCTCCTCTGCCTGCTGCACGCGGGACAATCCGCGCCAGGCG 120

Qy   121 CCGGGGCGCTCTTCACCCTGAACCTCACGTGCGGGGACCTGCTGTTGTGCACCGTGGTCAACATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCGGGCGCTCTTCACCCTGAACCTCACGTGCGGGGACCTGCTGTTGTGCACCGTGGTCAACATG 180

Db   181 CCGCTCACGCTGGGCCGCCGCGCGGCGCGCCGCGCGGGGGCGCGCAGCCGCGGCGCGGCGCGGCG 240

```
RESULT 2
LOCUS       CQ831051      1011 bp    DNA    linear   PAT 29-JUL-2004
DEFINITION  CQ831051 Sequence 1 from Patent EP1437595.
ACCESSION   CQ831051
VERSION     CQ831051.1  GI:5081178
KEYWORDS
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Goodnow,J.R., Mark,D.F., Martin,M.L. and Rosinski,J.A.
  TITLE     Orphan GPCR gene 115 for obesity indication
  JOURNAL   Patent: EP 1437595-A 1 14-JUL-2004;
            F. HOFFMANN-LA ROCHE AG (CH)
FEATURES             Location/Qualifiers
     source          1..1011
                     /organism="Mus musculus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10090"

ORIGIN

Query Match           99.7%;  Score 1011;  DB 6;  Length 1011;
Best Local Similarity 100.0%;  Pred. No. 1.7e-123;
Matches 1011;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

DNA sequence alignment (Qy = Query, Db = Database subject), spanning positions 1–1011.

```
RESULT 3

Db   961 TCTGGGCCTCACAGGCGACTCTTCACAGCCAGAGAACATTTCTGCCCGGTGTCTGAG 1011
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy       TCTGGGCCTCACAGGCGACTCTTCACAGCCAGAGAACATTTCTGCCCGGTGTCTGAG
```

LOCUS   HSA505757   7512 bp   mRNA   linear   PRI 05-SEP-2002
DEFINITION Homo sapiens mRNA for putative orphan G protein-coupled receptor 26
            (GPR26 gene).
ACCESSION   AJ505757
VERSION   AJ505757.1 GI:22293640
KEYWORDS   GPR26 gene; orphan G protein-coupled receptor 26.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Boulay,J.L., Labuhn,M., Jones,G., Maier,D. and Merlo,A.
  TITLE   The 10q25.3-26.1 gene encoding the orphan G protein-coupled
            receptor GPR26 is epigenetically silenced in human gliomas
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Jones,G., Boulay,J.L., Maier,D. and Merlo,A.
  TITLE   Sequence of the human homologue of the rat orphan G protein-coupled
            receptor GPR26 mRNA
  JOURNAL   Unpublished
REFERENCE   3 (bases 1 to 7512)
  AUTHORS   Boulay,J.L.
  TITLE   Direct Submission
  JOURNAL   Submitted (13-AUG-2002) Boulay J.L., Research, University Hospital
            of Basel, 46 Schanzenstrasse, CH-4031 Basel, SWITZERLAND
FEATURES   Location/Qualifiers
     source   1..7512
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="10q25-26"
                     /tissue_type="brain"
                     /dev_stage="foetus"
     gene   1..7512
                     /gene="GPR26"
     CDS   53..1066
                     /gene="GPR26"
                     /codon_start=1
                     /product="putative orphan G protein-coupled receptor 26"
                     /protein_id="CAD44281.1"
                     /db_xref="GI:22293641"
                     /db_xref="GOA:Q8NDV2"
                     /db_xref="UniProt/TrEMBL:Q8NDV2"
                     /translation="MNSWDAGLAGILVGTMGVSLLSNALVLLCLLHSADIRROAPALF
                     TLNLTCVDLLCTVVNMPLTLAGVVARQPQAGHDLCRLRAAFLDTFLAANSMLSMAALSI
                     DRWVAVVFPLSYRAKMRLRDAALMVAYTWLHALTFPAAIALSWLGFHOLYASCTLCS
                     RRPDERLRFAVFTGAFHALSFLLSFVVLLCCTYLKVLRVARFHCKRIDVITMQTIVLLV
                     DLHPSVRRCLEEQKRRQRATKKISTFIGTFIVCFAPYVITRLVELFSTVPIGSHWG
                     VLSKCLAYSKAASDPFVYSLLRHQYRKSCKEILNRLLHRRSIHSSGLTGDSHSQNILP
                     VSE"
```

ORIGIN

```
Query Match         99.7%;  Score 1010.8;  DB 9;  Length 7512;
Best Local Similarity   99.8%;  Pred. No. 1.1e-123;
Matches 1012; Conservative  0; Mismatches  2;  Indels  0;  Gaps  0;
```

```
Qy    1 ATGAACTCGTGGGACGGCGGCGGGCCCTGGCCGGGCGGGGGGGCTGGTGGGCACGATGGGCGGGTC 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   53 ATGAACTCGTGGGACGGCGGCGGGCCCTGGCCGGGCGGGGGGGCTGGTGGGCACGATGGGCGGGTC 112

Qy   61 CTGTCCAACGCGGCTGGTGCTGCTCTGCCTGCTCCACAGCGCGGACATCCGCCGCCAGGCG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  113 CTGTCCAACGCGCTGGTGCTGCTCTGCCTGCTCCACAGCGCGGACATCCGCCGCCAGGCG 172

Qy  121 CCGGCGCTCTTCACCCTGAACCTCACGTGCGTGGGACGGGGAACCTGCTGTGTGCAACATG 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  173 CCGGCGCTCTTCACCCTGAACCTCACGTGCGTGGGACGGGGAACCTGCTGTGTGCAACATG 232
```

```
Db  173 CCGGCGCTCTTCACCCTGAACCTCACGTGCGTGGGACGGGGAACCTGCTGTGTGCAACATG 232
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  233 CCGCTCACGCTGGCCGGCGTCGTGGCCCGGCAGCCGCAGGCGGGGCACGACCTGTGCCGC 292
Qy  181 CCGCTCACGCTGGCCGGCGTCGTGGCCCGGCAGCCGCAGGCGGGGCACGACCTGTGCCGC 240
        
Qy  241 CTGGCCGCCTTCCTCGACACCTTCCTGGCCGCCAACTCCATGCTCAGCATGGCCGCGCTC 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  293 CTGGCCGCCTTCCTCGACACCTTCCTGGCCGCCAACTCCATGCTCAGCATGGCCGCGCTC 352

Qy  301 AGCATCGACCGCTGGGTGGCCGTGGTCTTCCCGCTGAGCTACCGGGCCAAGATGCGCCTC 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  353 AGCATCGACCGCTGGGTGGCCGTGGTCTTCCCGCTGAGCTACCGGGCCAAGATGCGCCTC 412

Qy  361 CGCGACGCCGCGCTCATGGTGGCCTACACGTGGCTGCACGCCCTGACCTTCCCGGCCGCC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  413 CGCGACGCCGCGCTCATGGTGGCCTACACGTGGCTGCACGCCCTGACCTTCCCGGCCGCC 472

Qy  421 GGGCTCGCCCTGTCCTGGCTTGGGTTCCACCAGCTGTACGCCTCGTGCACGCTGTGCAGC 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  473 GGGCTCGCCCTGTCCTGGCTTGGGTTCCACCAGCTGTACGCCTCGTGCACGCTGTGCAGC 532

Qy  481 CGGCGGCCCAGACGAGGCCGCCTGCGCTTCGCCGTCTTCACTGGGGCCTTCCACGCTCAGC 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  533 CGGCGGCCCGGACGAGGCCGCCTGCGCTTCGCCGTCTTCACTGGGGCCTTCCACGCTCAGC 592

Qy  541 TTCCTGCTCTCCTTCGTGGTCTTGCTGTGCTGCACGTACCTCAAGGTGCTCAAGGTGGCC 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  593 TTCCTGCTCTCCTTCGTGGTCTTGCTGTGCTGCACGTACCTCAAGGTGCTCAAGGTGGCC 652

Qy  601 TTCCATTGCAAGCGCATCGACGTGATCACCATGCAGACGATCGTGCTGCTGGTGGACCTG 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  653 TTCCATTGCAAGCGCATCGACGTGATCACCATGCAGACGATCGTGCTGCTGGTGGACCTG 712

Qy  661 CACCCCAGTGTGCGGAGGTGCCTGGAGGAGCAGAAGCGGCGGCAGCGGGCCACC 720
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  713 CACCCCAGTGTGCGGAGGTGCCTGGAGGAGCAGAAGCGGCGGCAGCGGGCCACC 772

Qy  721 AAGAAGATCAGCACCTTCATAGGGACCTTCTCGGTGTGCTTCGCGCCTATGTGATCACC 780
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  773 AAGAAGATCAGCACCTTCATAGGGACCTTCTCGGTGTGCTTCGCGCCTATGTGATCACC 832

Qy  781 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTCCCACTGGGGGGTGCTGTCCAAG 840
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  833 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTCCCACTGGGGGGTGCTGTCCAAG 892

Qy  841 TGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGGGACACCAG 900
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  893 TGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGGGACACCAG 952

Qy  901 TACCGGCAAAAGCTGCAAGGAGAGATTCTGAACAGGCTCCTGCACAGGAGCCTCCATCCACTCC 960
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  953 TACCGGCAAAAGCTGCAAGGAGAGATTCTGAACAGGCTCCTGCACAGGAGCCTCCATCCACTCC 1012

Qy  961 TCTGGGCCTCACAGGCCACTCTCACAGCCAGAGAACATTCTGCGCGGTGTCTGAGTGA 1014
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1013 TCTGGGCCTCACAGGCGACTCTTCACAGCCAGAGAACATTTCTGCGCGGTGTCTGAGTGA 1066
```

RESULT 4

```
LOCUS       BD130531                2581 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  Ligand receptors and utilization thereof.
ACCESSION   BD130531
VERSION     BD130531.1  GI:23225476
KEYWORDS    JP 2002501083-A/1.
SOURCE      unidentified
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2581)
  AUTHORS   Glucksmann,A.M. and Robison,K.
  TITLE     Ligand receptors and utilization thereof
  JOURNAL   Patent: JP 2002501083-A 1 15-JAN-2002;
```

COMMENT   MILLENNIUM PHARMACEUTICALS INC
OS   Unidentified
PN   JP 200501083-A/1
PD   15-JAN-2002
PF   20-JAN-1999 JP 200528599
PR   26-JAN-1998 US 09/013634
PI   ALEXANDRA M GLUCKSMANN, KEITH ROBISON
PC   C07K14/705, C07K16/28, C12N5/09, C12N15/09, C12P21/02, C12Q1/68, PC
      G01N33/53,
PC   C12N5/00, C12N15/00
CC   Strandedness: Single;
CC   Topology: Linear;
CC   Ligand receptors and utilization thereof
FH   Key   Location/Qualifiers
FT   CDS   184..1194.

FEATURES   Location/Qualifiers
     source   1..2581
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"

ORIGIN

Query Match
Best Local Similarity 99.5%;   Score 1009.2; DB 6; Length 2581;
Matches 1011; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy     1 ATGAACTCGTGGGACGCGGGGACCGGGCCTACTCGGTGGGCACGAGGGTCTCGCTG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 ATGAACTCGTGGGACGCGGGGACCGGGCCTACTCGGTGGGCACGAGGGTCTCGCTG  243

Qy    61 CTGTCCAACGCGGCTGGTGCTGCTCCTCTGCACGGGGACATCCCGCCCGCCAGGCG  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   244 CTGTCCAACGCGGCTGGTGCTGCTCCTCTGCACGGGGACATCCCGCCCGCCAGGCG  303

Qy   121 CCGGGGCGCTCTTCACCCTGAACCTCACGTGCGGGGAGAACCTGCTGTGTCAACATG  180
         |||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db   304 CCGGGGCGCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGCTGTGTCAACATG   363

Qy   181 CCGCTCACGCTGGCGCCGGGCGTCGTGGGGCGGAGCAGCCGGGACGCGCCGTTGCGCGC  240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   364 CCGCTCACGCTGGCGCCGGGCGTCGTGGGGCGGAGCAGCCGGGACGCGCCGTTGCGCGC  423

Qy   241 CTGGGTGCCGTTCCTGACACACCTTCCTGGGCTGCCAACTCCATGCTCAGCATGGCCCGC  300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   424 CTGGGTGCCGTTCCTGACACACCTTCCTGGGCTGCCAACTCCATGCTCAGCATGGCCCGC  483

Qy   301 AGCATCGACCCGGGCTGGGTGGTGGTTCTTCCCGGCTGAGCTACCGGGCCAAGATGCCCTC  360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   484 AGCATCGACCCGGGCTGGGTGGTGGTTCTTCCCGGCTGAGCTACCGGGCCAAGATGCCCTC  543

Qy   361 CGCGACGCGGCGGCTCATGGTGGCCTACACACGTGGCTGCACGGCGGCTCACCTTCCCAGCCGCC  420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   544 CGCGACGCGGCGGCTCATGGTGGCCTACACACGTGGCTGCACGGCGGCTCACCTTCCCAGCCGCC  603

Qy   421 GCGCTCGCCCCTGTCCTGGCTGGCTTCCACCAGCTGTACGCCTCGTCGTGTGCGCGC  480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   604 GCGCTCGCCCCTGTCCTGGCTGGCTTCCACCAGCTGTACGCCTCGTCGTGTGCGCGC  663

Qy   481 CGGGCGCGCCAGACGACGAGCGCCTCGCGCTTCGCGCGCTTCACTCGGCTCTCAGC  540
         |||||||||| ||||| ||||||||||||| ||||||||||||||||||||||||
Db   664 CGGGCGCGCCGGACGAGCGCCCTCGCGGTGGCGCCTTCACTGCGGCTCTCAGC   723

Qy   541 TTCCTGCTCTCCTTCGTGCTTCGTGCTGCTGCTCCTCAAGGTGCTCAAGGTGCCCGC  600
         |||||||||||||||||||| |||||||||||| ||||||||||||||||||||||
Db   724 TTCCTGCTCTCCTTCGTGCTGCTGCTGCTGCTGCTCAAGGTGCTCAAGGTGCCGCTG  783

Qy   601 TTCCATTGCAAGGCATCGAGCGTGATCACCATGCAGACGCTGGTGCTGCTGGTGGACCTG  660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   784 TTCCATTGCAAGGCATCGAGCGTGATCACCATGCAGACGCTGGTGCTGCTGGTGGACCTG  843

Qy   661 CACCCCAGTGTGCGGGGAAACGCTGTCTGGAGGAGGAGCAGCACGAGCCGCC  720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   844 CACCCCAGTGTGCGGGGAAACGCTGTCTGGAGGAGGAGCAGCACGAGCCGCC  903
```

Qy   721 AAGAGAGATCAGCACCTTCATAGGGACGACCTTCCTTGTTGTTGTTGCTTGCGCGCCCTATGTGATCACC 780
```
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   904 AAGAGAGATCAGCACCTTCATAGGGACGACCTTCCTTGTTGTTGTTGCTTGCGCGCCCTATGTGATCACC 963

Qy   781 AGGCTAGTGGAGCGCTCTTCTCCACCGTGCCCATCGCTGCCCATCCCACTGGGGGTCGTCCAAG 840
```
         |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db   964 AGGCTAGTGGAGCTCTTCTCCACCGTGCCCATCGCTGCCCATCCCACTGGGGGTCGTCCAAG 1023

Qy   841 TGCTTGGCGTACAGCAGAAGGCCCATCCGACCCTTTGTGTACTCCTTACTGGCGACACCAG 900
```
         ||||||||||||||||||||||| |||||||||| |||||||||||||||| ||||||||||
```
Db 1024 TGCTTGGCGTACAGCAGAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGCGACACCAG 1083

Qy   901 TACCGGCAAAAGCTGCAAAGAGAGATTCTGAACACGGGTGCTCCTGCACAGAGAGGCTCCACTCC 960
```
         |||||||||||||||||||| ||||||||||||| |||||||||||||||||||||||||||||
```
Db 1084 TACCGGCAAAAGCTGCAAAGGAGGATTCTGAACAGGGTGCTCCTGCACAGAGAGGCTCCACTCC 1143

Qy   961 TCTGGCCTCACAGGGCGACTCTCACACCAGGACAATTCTGCCGGTGTCTGACTGA 1014
```
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
```
Db 1144 TCTGGCCTCACAGGGCGACTCTCACACCAGGACAATTCTGCCGGTGTCTGACTGA 1197

RESULT 5
BD130532
LOCUS   BD130532   1011 bp   DNA   linear   PAT 18-SEP-2002
DEFINITION Ligand receptors and utilization thereof.
ACCESSION   BD130532
VERSION   BD130532.1 GI:23225477
KEYWORDS   JP 200501083-A/2.
SOURCE   unidentified
  ORGANISM unidentified
            unclassified.
REFERENCE   1 (bases 1 to 1011)
  AUTHORS   Glucksmann,A.M. and Robison,K.
  TITLE   Ligand receptors and utilization thereof
  JOURNAL   Patent: JP 200501083-A 2 15-JAN-2002;
            MILLENNIUM PHARMACEUTICALS INC
COMMENT   OS   Unidentified
PN   JP 200501083-A/2
PD   15-JAN-2002
PF   20-JAN-1999 JP 200528599
PR   26-JAN-1998 US 09/013634
PI   ALEXANDRA M GLUCKSMANN, KEITH ROBISON
PC   C07K14/705, C07K16/28, C12N5/10, C12N15/09, C12P21/02, C12Q1/68, PC
      G01N33/53,
PC   C12N5/00, C12N15/00

FEATURES   Location/Qualifiers
     source   1..1011.
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
     CDS   1..1011.

ORIGIN

Query Match
Best Local Similarity 99.2%;   Score 1006.2; DB 6; Length 1011;
Matches 1008; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy     1 ATGAACTCGTGGGACGCGGGGACCGGGCCTACTCGGTGGGCACGAGGGTCTCGCTG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGAACTCGTGGGACGCGGGGACCGGGCCTACTCGGTGGGCACGAGGGTCTCGCTG   60

Qy    61 CTGTCCAACGCGGCTGGTGCTGCTCCTCTGCACGGGGACATCCCGCCCGCCAGGCG  120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTGTCCAACGCGGCTGGTGCTGCTCCTCTGCACGGGGACATCCCGCCCGCCAGGCG  120

Qy   121 CCGGGCGCTCTTCACCCTGAACCTCACCTGCGGGGAAACCTGCTGTGTCAACATG  180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||

```

```
Qy        181 CCGCTCACGCTGGCCGGCGTCGTGGCGCAGCGGCAGCCGGCGGGCGACCGCCTGTGCCGC 240
              |||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db        181 CCGCTCACGCTGGCCGGCGTCGTGGCGCGGCGGCAGCCGGCGGGCGACCGCCTGTGCCGC 240

Qy        241 CTGGCTGCCTTCCTCGACACCTTCCTGGCTGCCAACTCCATGCTCAGCATGGCCGCGCTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGCTGCCTTCCTCGACACCTTCCTGGCTGCCAACTCCATGCTCAGCATGGCCGCGCTC 300

Qy        301 AGCATCGACCGCTGGGTGGCCGTGGTCTTCCCGCTGAGCTACCGGGCCAAGATGCGCCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCATCGACCGCTGGGTGGCCGTGGTCTTCCCGCTGAGCTACCGGGCCAAGATGCGCCTC 360

Qy        361 CGCGACGCGGCGCTCATGGTGGCCTACACGTGGCTGCACGCGCTCACCTTCCCAGCCGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCGACGCGGCGCTCATGGTGGCCTACACGTGGCTGCACGCGCTCACCTTCCCAGCCGCC 420

Qy        421 GCGCTCGCCCTGTCCTGGCTCGGCTTCCACCAGCTGTACGCCTCGTGCACGCTGTGCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GCGCTCGCCCTGTCCTGGCTCGGCTTCCACCAGCTGTACGCCTCGTGCACGCTGTGCAGC 480

Qy        481 CGGCGGCCAGACGAGCGCCTGCGCTTCGCCGTCTTCACTGGCGCCTTCCACGCTCTCAGC 540
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CGGCGGCCGGACGAGCGCCTGCGCTTCGCCGTCTTCACTGGCGCCTTCCACGCTCTCAGC 540

Qy        541 TTCCTGCTCTCCTTCGTCGTGCTCTGCTGCACGTACCTCAAGGTGCTCAAGGTGGCCCGC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTCCTGCTCTCCTTCGTCGTGCTCTGCTGCACGTACCTCAAGGTGCTCAAGGTGGCCCGC 600

Qy        601 TTCCATTGCAAGCGCATCGACGTGATCACCATGCAGACGCTGGTGCTGCTGGTGGACCTG 660
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        601 TTCCATTGCAAGCGCATCGACGTGATCACCATGCAGACGCTCGTGCTGCTGGTGGACCTG 660

Qy        661 CACCCCAGTGTGCGGGAACGCTGTCTGGAGGAGCAGAAGCGGAGGCGACAGCGAGCCACC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CACCCCAGTGTGCGGGAACGCTGTCTGGAGGAGCAGAAGCGGAGGCGACAGCGAGCCACC 720

Qy        721 AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGCTTCGCGCCCTATGTGATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGCTTCGCGCCCTATGTGATCACC 780

Qy        781 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTCCCACTGGGGGGTGCTGTCCAAG 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTCCCACTGGGGGGTGCTGTCCAAG 840

Qy        841 TGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGCGACACCAG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGCGACACCAG 900

Qy        901 TACCGCAAAAGCTGCAAGGAGATTCTGAACAGGCTCCTGCACAGACGCTCCATCCACTCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACCGCAAAAGCTGCAAGGAGATTCTGAACAGGCTCCTGCACAGACGCTCCATCCACTCC 960

Qy        961 TCTGGCCTCACAGGCGACTCTCACAGCCAGAACATTCTGCCGGTGTCTGAG 1011
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTGGCCTCACAGGCGACTCTCACAGCCAGAACATTCTGCCGGTGTCTGAG 1011

RESULT 6
CQ729494
LOCUS       CQ729494                 969 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 15428 from Patent WO02068579.
ACCESSION   CQ729494
VERSION     CQ729494.1  GI:42300819
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 15428 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .969
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             95.4%;  Score 967.4;  DB 6;  Length 969;
  Best Local Similarity   99.9%;  Pred. No. 9.2e-118;
  Matches  968;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         46 ATGGGCGTCTCGCTGCTGTCCAACGCGCTGGTGCTGCTCTGCCTGCTGCACAGCGCGGAC 105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGCGTCTCGCTGCTGTCCAACGCGCTGGTGCTGCTCTGCCTGCTGCACAGCGCGGAC 60

Qy        106 ATCCGCCGCCAGGCGCCGGCGCTCTTCACCCTGAACCTCACGTGCGGGAACCTGCTGTGC 165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ATCCGCCGCCAGGCGCCGGCGCTCTTCACCCTGAACCTCACGTGCGGGAACCTGCTGTGC 120

Qy        166 ACCGTGGTCAACATGCCGCTCACGCTGGCCGGCGTCGTGGCGCAGCGGCAGCCGGCGGGC 225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGTGGTCAACATGCCGCTCACGCTGGCCGGCGTCGTGGCGCAGCGGCAGCCGGCGGGC 180

Qy        226 GACCGCCTGTGCCGCCTGGCTGCCTTCCTCGACACCTTCCTGGCTGCCAACTCCATGCTC 285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GACCGCCTGTGCCGCCTGGCTGCCTTCCTCGACACCTTCCTGGCTGCCAACTCCATGCTC 240

Qy        286 AGCATGGCCGCGCTCAGCATCGACCGCTGGGTGGCCGTGGTCTTCCCGCTGAGCTACCGG 345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCATGGCCGCGCTCAGCATCGACCGCTGGGTGGCCGTGGTCTTCCCGCTGAGCTACCGG 300

Qy        346 GCCAAGATGCGCCTCCGCGACGCGGCGCTCATGGTGGCCTACACGTGGCTGCACGCGCTC 405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCCAAGATGCGCCTCCGCGACGCGGCGCTCATGGTGGCCTACACGTGGCTGCACGCGCTC 360

Qy        406 ACCTTCCCAGCCGCCGCGCTCGCCCTGTCCTGGCTCGGCTTCCACCAGCTGTACGCCTCG 465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACCTTCCCAGCCGCCGCGCTCGCCCTGTCCTGGCTCGGCTTCCACCAGCTGTACGCCTCG 420

Qy        466 TGCACGCTGTGCAGCCGGCGGCCAGACGAGCGCCTGCGCTTCGCCGTCTTCACTGGCGCC 525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGCACGCTGTGCAGCCGGCGGCCAGACGAGCGCCTGCGCTTCGCCGTCTTCACTGGCGCC 480

Qy        526 TTCCACGCTCTCAGCTTCCTGCTCTCCTTCGTCGTGCTCTGCTGCACGTACCTCAAGGTG 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTCCACGCTCTCAGCTTCCTGCTCTCCTTCGTCGTGCTCTGCTGCACGTACCTCAAGGTG 540

Qy        586 CTCAAGGTGGCCCGCTTCCATTGCAAGCGCATCGACGTGATCACCATGCAGACGCTGGTG 645
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        541 CTCAAGGTGGCCCGCTTCCATTGCAAGCGCATCGACGTGATCACCATGCAGACGCTCGTG 600

Qy        646 CTGCTGGTGGACCTGCACCCCAGTGTGCGGGAACGCTGTCTGGAGGAGCAGAAGCGGAGG 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTGCTGGTGGACCTGCACCCCAGTGTGCGGGAACGCTGTCTGGAGGAGCAGAAGCGGAGG 660

Qy        706 CGACAGCGAGCCACCAAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGCTTCGCG 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CGACAGCGAGCCACCAAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTGCTTCGCG 720

Qy        766 CCCTATGTGATCACCAGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTCCCACTGG 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CCCTATGTGATCACCAGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTCCCACTGG 780

Qy        826 GGGGTGCTGTCCAAGTGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCC 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 GGGGTGCTGTCCAAGTGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCC 840

Qy        886 TTACTGCGACACCAGTACCGCAAAAGCTGCAAGGAGATTCTGAACAGGCTCCTGCACAGA 945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTACTGCGACACCAGTACCGCAAAAGCTGCAAGGAGATTCTGAACAGGCTCCTGCACAGA 900
```

RESULT 7

```
LOCUS       CQ831053                1172 bp    DNA     linear   PAT 29-JUL-2004
DEFINITION  Sequence 3 from Patent EP1437595.
ACCESSION   CQ831053
VERSION     CQ831053.1  GI:5083179
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Goodnow,J.R., Mark,D.F., Martin,M.L. and Rosinski,J.A.
  TITLE     Orphan GPCR gene 115 for obesity indication
  JOURNAL   Patent: EP 1437595-A 3 14-JUL-2004;
            F. HOFFMANN-LA ROCHE AG (CH)
FEATURES             Location/Qualifiers
     source          1..1172
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN
```

Query Match 85.6%; Score 868.4; DB 6; Length 1172;
Best Local Similarity 91.0%;
Matches 923; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```
Qy      1 ATGAACTCGTGGGACGGCGAGCGGCGGCCTGGCGGGCGCCGCTACTGGTGGGCGCACGATAGGGCGTCTTCCGCTG   60
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     27 ATGAACTCGTGGGACGGCGAGCGGCGGCCTGGCGGGCGCCGCTGCTGGTGGGCACNAATCCGGGCGTGTCGGCTG   86

Qy     61 CTGTCCAAGGCGCGTGTGCTGCTGCTGCTGCTGCTGCACAGCCGCGACATCCCCCGCCCAGGCG   120
          |||||||||||||||||||||||||||||||||||||||||||||
Db     87 CTGTCCAACGCGGGCGTGGGTGCTGCGCTCCTGCACCGCGGCTGACATCCGCCGCCCAGGCG   146

Qy    121 CCGGCGCGCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGTGCTGGTGTGGGTGAACATG   180
          ||||||||||||||||||||||||||||||||||||||||||
Db    147 CCGGCGCGCTCTTCACTCTCAACCTCACGTGTGGTGTACCCGTGGTCTGGGTGGTGAACATG   206

Qy    181 CCGCTCACGCTGGCGCCGGGCCGTGCGCTGGGCCGAGCCGACCTGCCCTGTGCCCGC   240
          |||||||||||||||||||||||||||||||||||||||||
Db    207 CCACTAACACTGGCCGCGGGCGTGGTGCGACACGGCAGCGCCGACCCCTGTGCCCGC   266

Qy    241 CTGGCGGCGTTCCTGCTGGACACACCTTTCCTGGGCGTGCCAACTCCATGCTCAGTTATGGGCCGGCTC   300
          ||||||||||||||||||||||||||||||||||||||||||||
Db    267 CTGGCGCCGTTCCTGCGACACCTTTCTGGGCGCGCCAACTCCATGCTCAGTGCTCAGTTATGGGCCGGCTC   326

Qy    301 AGCATCGACGCGGGCTGGGCGCCGTGGCTGCCGAGCTACACCGGCGCCAAGATGCCCCTC   360
          |||||||||||||||||||||||||||||||||||||||||
Db    327 AGCATCGACGCGGGTGGTGTCTTCCCGCTACCGTGCCCAAGATGCCGCCTC   386

Qy    361 CGGCGACGCGGCGGCGCTCATGGTGGCCTACACGGTGGCTGCAGCGCGCACAGTTCCCAGCCGCC   420
          |||||||||||||||||||||||||||||||||||
Db    387 CGAGATGCCGGCCTTCATGGTGGTGGCCTACACGCCGGCGCTCAGCACCGTCCAGCCCCACC   446

Qy    421 GCGTTCGACCCTGTTCGGCTTCGGCTTCCACCAGGCTGTACGGCCTGTGCAGC   480
          ||||||||||||||||||||||||||||||||||||||||
Db    447 GCGCTTCGCCCTGTTCCTGGGCTTCTATGCCTCGTTGCTGCGCCTGCGCAGC   506

Qy    481 CGGCGCGGCCCAGACGAGCGCGCTCGCGCCTTCGCTTCGCGCTTTGCGTGTTCTTCACTGGGCTCTCAGC   540
          ||||||||||||||||||||||||||||||||||||||||
Db    507 CGGCGGCCCGCAGAGGAGCGCCTGCGCGCCTTCGCCTTTAGC   566
```

RESULT 8

```
LOCUS       AF208288                1172 bp    mRNA    linear   ROD 13-MAR-2000
DEFINITION  Rattus norvegicus orphan G protein-coupled receptor GPR26 (GPR26)
            mRNA, complete cds.
ACCESSION   AF208288
VERSION     AF208288.1  GI:6644327
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1172)
  AUTHORS   Lee,D.K., Lynch,K.R., Nguyen,T., Im,D.S., Cheng,R., Saldivia,V.R.,
            Liu,Y., Liu,I.S., Heng,H.H., Seeman,P., George,S.R., O'Dowd,B.F.
            and Marchese,A.
  TITLE     Cloning and characterization of additional members of the G
            protein-coupled receptor family
  JOURNAL   Biochim. Biophys. Acta 1490 (3), 311-323 (2000)
  MEDLINE   20149852
   PUBMED   10684976
REFERENCE   2  (bases 1 to 1172)
  AUTHORS   Lee,D.K., Lynch,K.R., Nguyen,T., Im,D.-S., Cheng,R., Saldivia,V.R.,
            Liu,Y., Liu,I.S.C., Heng,H.H.Q., Seeman,P., George,S.R.,
            O'Dowd,B.F. and Marchese,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-NOV-1999) Pharmacology, Univ. Virginia, 1300
            Jefferson Park Avenue, Charlottesville, VA 22908, USA
FEATURES             Location/Qualifiers
     source          1..1172
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /chromosome="10"
                     /map="10q26.2-q26.3"
     gene            1..1172
                     /gene="GPR26"
     CDS             27..1040
```

Query Match ... Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```
Qy    541 TTCCTGGCTCTCCTTCGTTCGTTCGTGCTTCGTGCTGCACGTACCTCAAGGTGCTCAAGGTGGCCGC   600
          |||||||||||||||||||||||||||||||||||||||||
Db    567 TTCCTGGCTCTCCTTCCATCGTTCGTCGCTGCTTCACGTCGTTCGCTTCAAGGTTGCTCAAGGTGGCCCGT   626

Qy    601 TTCCATTGCAAGCGCCATCCAGCGTGATCACCATGCAGAGAGCGACACCCTCGAAGGTGGGACCTG   660
          |||||||||||||||||||||||||||||||||||||||||
Db    627 TTCCAACTGCAAGCGCATCGACGTGATCACCATGCAGAGGAGCCGTGTTAGTGGACATC   686

Qy    661 CACCCCAGTGTGGGGAACCGTGTCGTGGAGGAGGAGAAGCGAGGCGAGGAGCCACC   720
          ||||||||||||||||||||||||||||||||||||||||
Db    687 CATCCCAGTGTGCAGGGAACGAGTCTGAGGAACAGAGAGCGGGAGGGAGGAGGAGGCGAGGAGCCACC   746

Qy    721 AAGAAGATCAGCACCTTCAATAGGGACACTTCCTTGTGTTGCTCCGCGCCTTATGTGGATCAGC   780
          ||||||||||||||||||||||||||||||||||||||||
Db    747 AAGAGAGATCAGCACCTTCAATAGGGACACTTCCTTCGTCGTTGGCTCGCCCCTATGTGGATTACC   806

Qy    781 AGGGCTAGTGAGGTCTTCTCCACGGTGCCCATCCGCGTCCCCACTCGGGGGGTGGTGCTTCCTCCAAG   840
          ||||||||||||||||||||||||||||||||||||
Db    807 AGGGCTGGTGAACTCTTCTCCACAGCACCCCATTGACTCACACTGGGGTGTGGTGGTTCCAAG   866

Qy    841 TGCTTGGCGGTACAGCAAGGCGGCCATCGACCCCCCTTTGTGGTAGTGGTCTTCCTACTGGGACACCAG   900
          |||||||||||||||||||||||||||||||||||||||||
Db    867 TGCTTGGCCTACAGCAAGGCGGCTTCTTGACCCCCTTCGTGTGTACTCCTGTTGGTCCTGCGGACACCAG   926

Qy    901 TACCCGCCAAAGCTGCAAGGAGATTCTGAACACGGTGCCGCACAGAGACGGCTCCATCCACTCC   960
          |||||||||||||||||||||||||||||||||||||||||
Db    927 TACCCGCGAGGAGCTGCAAGGAGCTTCTGAACACAGGATCTTCAACAGAGACGCTCCATCCATTCACTCT   986

Qy    961 TCTGGCCTCACAGGCGACTCCACAGCCCAGACAGCATTCTGCGGGTTCTGAGTGA   1014
          ||||||||||||||||||||||||||||||||||||
Db    987 GTGGGCCTCACAGGTGACTCTCACAGCCAGAACATTCTGCCCAGTGTCTCGGGAATGA   1040
```

/gene="GPR26"
/codon_start=1
/product="orphan G protein-coupled receptor GPR26"
/protein_id="AAF21012.1"
/db_xref="GI:6643328"
/translation="MNSWDAGLLAGLLVGTIGVSLLSNGIVLLCLLHSADIRRQAPALF
TLNLTCGNLLCTVVNMPLTLAGVVAQRQPAGDRLCRLAAFLDTFLAANSMLSWAALSI
DRWVAVVFPLSYRAKMRLRDAAFMVAYTWLHALTPAATALALSWLGFHQLYASCTLCS
RRPDERLRPAVFTSAFHALSFLLSFIVLCFTYLKVLKVARFHCKRIDIVTMQTLLVLLV
DIHPSVRRCLEQGKRRQROATKKISTFICTFIVCFAPVVITRLVELPSTAPIDSHWG
VLSKCLAYSKAASDPFVYSLLRHQYRRSCKELLNRIFNNRSIHSVGLTSDSHSQNILP
VSE"

ORIGIN
Best Local Similarity   85.6%;   Score 868.4; DB 10; Length 1172;
Matches 923; Conservative 91.0%; Pred. No. 8.4e-105;
         0; Mismatches 91; Indels   0; Gaps   0;

```
Qy     1 ATGAACTCGTGGGACGCGGGACGCGGCGGGCTGGTGGCGGGCCTACTGGTGGGGCACACGGGCGGGGGTGTCGCTG      60
Db    27 ATGAACTCGTGGGACGCGGGACGCGGCGGGCTGCTGGCGGGCCTACTGGTGGGGCACAATGCGGCGGTCGCTG      86

Qy    61 CTGTCCAACGCGCTGGGTGCTGTCTCTGCCTGCTGCAACAGCGGACATCCCGCGCCAGGCCG      120
Db    87 CTGTCCAACGCGCTGGGTGCTGCTGCTGCTGCTGCTGCAACAGCGGACATCCCGCGCCAGGCCG      146

Qy   121 CCGGCGGGCCTCTTCACCCTGAACCTCACGTGCGGGAACCTGCTGTGCACCGTGGTCAACATG      180
Db   147 CCGGCGGGCTCTTCACTCTCAACCTCACGTGTGGCACATCCCGTGTACCGTGGTCAACATG      206

Qy   181 CCGGCTCACGCTGGGGCCGGGGTCGTGGCGCAGCGGCAGCCGGCCGCTGGGTCAACTG      240
Db   207 CCACTAACACTGGGCGCCGGGGTCGTGGCACAACGGCAGCCGGCCGCTGTGCCGC      266

Qy   241 CTGGCGTGCCTTCCTCGACACCTTCCTGGCGTGCCAACTCCATGCTCAGGTCGACTG      300
Db   267 CTGCGCCGGCCCTTCCTCGACACCTTCCTGGCGGCCAACTCCATGCTCAGTATGGCCGCTC      326

Qy   301 AGCATCGACCGGGCTGGGTGGCCGTGGTTCCCGCTGAGCTACCGGGCCAAGATGCGGCTC      360
Db   327 AGCATCGACCGGGCTGGGTGGCTGTGGTCTTCCCGCTGAGCTACCGAGATGCGGCCTC      386

Qy   361 CGGCACGCGGCGCGCTCATGGTGGCCTACACGTGGCTGCACGCGCTCACCCCCAGCGCGCC      420
Db   387 CGAGATGCGCGCCTTCATGGTGGCCTACACGTGGCTCACCTTCCCGGCCACC      446

Qy   421 GCGCTGCGCCCTGTCCTGGGCGGGCTGTACGCGCTGGTGCACGCCAGC      480
Db   447 GCGCTGCGCCCTGTCCTGGGCTTCCACCAGCTGTATGCCTCGTGCACGCTGTGCAGC      506

Qy   481 CGGCGGCCCAGACAGAGGCCGCTGCGCCTGGCGTTCGCGGCGCCTTCGCCGTTCAGC      540
Db   507 CGGCGGCCCGACGAGAGGCTGCGCCCTGCGGTTCGTGTCTTCACAGCCTTCCATGCGTTAGC      566

Qy   541 TTCCTGCTCTCCTTCGTCGTTCTGCTTCGTGCTCAAGGTGCTCAAGGTGGCCCCGC      600
Db   567 TTCCTGCTCTCCTTCATCGTTCTGCTCGTTACCTCAAGGTGCTCAAGGTGGCCCCGC      626

Qy   601 TTCCATTGCAAGCCATCAGCGTGATCACCATGCAGACCCTGGTCGTGCTGGTGGAAGACCTG      660
Db   627 TTCCACTGCAAGCGCATCGATCACCATGCAGACCCTCGTGTTAGTGGACATC      686

Qy   661 CACCCCAGTGTGCGGGAGCTGTCTGGGAGGAGCAGGAGCGGACGAGCCACC      720
Db   687 CATCCCAGTGTGGAGCGGTGCCTGGAGGGAAGCGGCGGCGACCGGAGCCACC      746

Qy   721 AAGAGAGATCAGCACACCTTCATAGGGAACCTTCCTGTTGTGTTGTTCGGCGATCACC      780
Db   747 AAGAAGATCAGCACACCTTCATAGGGACCTTCCTCGTTGCTTTGCACCCTATGTGATTACC      806

Qy   781 AGGCTAGTGGGAGCTCTTCTCCACAGCACCCATTGGTGCACCCATGGACTCACACTGGGG      840
Db   807 AGGCTGGTCGGAACTCTTCTCCACAGCACCCATTGACTCACACTGGGGGTGGTGTCCAAG      866
```

```
Qy   841 TGCTTGGCGTACAGCAAGGCGCCGCGATCCGACCCCTTTGTGTACTCCTTACTGCGACACCAG      900
Db   867 TGCTTGGCCTACAGCAAGGCCGCTGCTTGACCCCCTTCGTGTACTCCTTGCTGCGACACCAG      926

Qy   901 TACCGCAAAAGCTGCAAGGAGATTCTGAACAGGATCTTCTGCGACAGAGCGCTCCATCCACTCC      960
Db   927 TACCGCAGGAGCTGCAAGGAGCTTCTGAACAGGATCTTCAACAGAGACCGCTCCATTCACTCT      986

Qy   961 TCTGGCCTCACAGGCGACTCTCACAGCCGACCAGAACAATTCTGCCGGTGTGTCTGAGTGA      1014
Db   987 GTGGGGCTCACAGGTGACTCTCACAGCCAGAGAACAATTCTGCCAGTGTCGGGAATGA      1040
```

Query Match   65.9%;   Score 668.6; DB 9;   Length 196808;
Best Local Similarity   98.7%;   Pred. No. 3.4e-79;
Matches 674; Conservative   0; Mismatches   9;   Indels   0;   Gaps   0;

```
Qy       1 ATGAACTCGTGGGACGCGGGACGCGGCGGGCTGGTGGCGGGCCTACTGGTGGGGCACACGGG      60
Db  160387 ATGAACTCGTGGGACGCGGGACGCGGCGGGCTGCTGGCGGGCCTACTGGTGGGGCACACGG      160328

Qy      61 CTGTCCAAGCGCGCTGGGTGCTGTCTGTGCCTGCTGCTGCTGCCCACACACAGCAGGGCTG      120
Db  160327 CTGTCCAACGCGCTGGGTGCTGTCTGTGCTGCTGCCTTGCTGCACAGCGCGGACATCCCGCCGCAGGCG      160268
```

```
Qy    121 CCGGCGCCTTCACCCTGAACCTCAGTGCGTCGGGGACCTCGGTGGTCAACATG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 160267 CCGGCGCCTTCACCCTGAACCTCAGTGCGTCGGGGAACCTCGCTGGTCAACATG 160208

Qy    181 CCGGCTCACCCTGGCCGCGGGCGGCCGCAGCGAGCGGAGCGGCAGGCCCTGTGCCGC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db 160207 CCGGCTCACCCTGGCCGCGGTCGTGGGCGGCAGCGGGCGACCGGCCCTGTGCCGC 160148

Qy    241 CTGGCGCTGCCTTCCTCGACACTTCCGTGCGTCCAACTCAGCATGTCGGCCGCGCTC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 160147 CTGGCGCTGCCTTCCTCGACAACTTCCGTGCGTCCAACTCAGCATGGCCGCGCTC 160088

Qy    301 AGCATCGACCGGCTGGGTGGTCTTGGCCCGTGGCGTGGCGACCTTCCAGCAGGTGCTC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 160087 AGCATCGACCGGCTGGGTGGGTCTTCCCCGTGGCGTGGCGACCTTCCAAGATGCGCTC 160028

Qy    361 CGGCCGACGCGCGGGGCTTCATGTGCGCGTCACCTTCCCAGCCGCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||
Db 160027 CGGCCGACGCGCGGGGCTCATGTGGTCGTCGTCACCTTCCCAGCCGCC 159968

Qy    421 GCGCTCGCCCCTGTCCTGGCGTCTCTCGCCGCCGTGCTGCCAGCCTGCCTCAGC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 159967 GCGCTCGCCCAGCCGCCGTTCTTCACGTCCTGGCGGTCCCACCCTGTGCCAGC 159908

Qy    481 CGGCGCCGCCAGAGAGGGCGCGCTTCGCCGCGTTCTTCACGGCCTCTCAGC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 159907 CGGCGCCGCCAGAGAGGGCGCGCTCGCCCCTTCCACGGCCTCTCAGC 159848

Qy    541 TTCCTGCTCTCCCTTCGTCGTGCGTGCGGTGGGTGCCCGGC 600
              |||||||||||||||||||||||||||||||||||||||||
Db 159847 TTCCTGCTCTCCCTTCGTCGTGCGTGCGGTGGGGTGCCCGGC 159788

Qy    601 TTCAATTGCAAGCGCCATCGAGCGTGATCACCATGCAGAGCGCTGGGACCTG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 159787 TTCCAATTGCAAGCGCCATCGAGCGTGATCACCATGCAGAGCGCTGGGACCTG 159728

Qy    661 CACCCCAGTGTGCGGGAACCCTG 683
              ||||||||||||||||||||||
Db 159727 CACCCCAGTGTGCGAGCCGCAG 159705
```

```
RESULT 10
LOCUS        AC105062          162502 bp    DNA     linear   ROD 26-JUN-2004
DEFINITION   Mus musculus chromosome 7, clone RP24-528E17, complete sequence.
ACCESSION    AC105062
VERSION      AC105062.13  GI:49258241
KEYWORDS     HTG.
SOURCE       Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1  (bases 1 to 162502)
  AUTHORS    Birren,B., Nusbaum,C. and Lander,E.
  TITLE      Mus musculus chromosome 7, clone RP24-528E17
  JOURNAL    Unpublished
REFERENCE    2  (bases 1 to 162502)
  AUTHORS    Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
             Anderson,M., Anderson,S., Barna,N., Bastien,V., Bloom,T.,
             Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
             Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
             Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
             Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
             Ginde,S., Grd,S., Goyette,M., Graham,L., Grand-Pierre,N.,
             Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
             Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
             Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
             Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
             McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
             Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
             Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,N.,
             Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
             Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
             Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
             Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
             Strauss,N., Subramanian,S., Arachchi,H.M., Barna,N., Bastien,V.,
             Topham,K., Travis,M., Travis,N., Trigilio,J., Tesfaye,S., Theodore,J.,
             Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Vassiliev,H.,
             Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE      Direct Submission
  JOURNAL    Submitted (23-DEC-2001) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    3  (bases 1 to 162502)
  AUTHORS    Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
             Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
             Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
             Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
             DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
             Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
             Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
             Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
             Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
             Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
             Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
             McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
             Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
             O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
             Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
             Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
             Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
             Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
             Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
             Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
             Zimmer,A. and Zody,M.
  TITLE      Direct Submission
  JOURNAL    Submitted (22-APR-2004) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE    4  (bases 1 to 162502)
  AUTHORS    Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
             Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
             Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
             Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,
             DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
             Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D.,
             Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
             Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
             Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
             Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
             Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C.,
             McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
             Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
             O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
             Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
             Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
             Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
             Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
             Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
             Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
             Zimmer,A. and Zody,M.
  TITLE      Direct Submission
  JOURNAL    Submitted (26-JUN-2004) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT      On Jun 26, 2004 this sequence version replaced gi:46485957.
             All repeats were identified using RepeatMasker:
             http://ftp.genome.washington.edu/RM/RepeatMasker.html

             --------------- Genome Center
             Center: Whitehead Institute/MIT Center for Genome Research
             Center code: WIBR
             Web site: http://www-seq.wi.mit.edu
             Contact: sequence_submissions@broad.mit.edu
             --------------- Project Information
             Center project name: L17510
             Center clone name: 528_E_17
             ---------------
FEATURES             Location/Qualifiers
```

```
source          1..162502
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /chromosome="7"
                /map="7"
                /clone="RP24-528E17"
                /clone_lib="RPCI-24 Male Mouse BAC"
misc_feature    1..4
                /note="clone_boundary
                clone_end:SP6
                site:MboI"
repeat_region   1065..1170
                /rpt_family="B1F"
repeat_region   1204..1255
                /rpt_family="B1F"
repeat_region   1295..1872
                /rpt_family="(TAAG)n"
repeat_region   1873..1924
                /rpt_family="Lx8"
repeat_region   1925..2211
                /rpt_family="CT-rich"
repeat_region   1925..2211
                /rpt_family="(CA)n"
repeat_region   2259..2690
                /rpt_family="Lx8"
repeat_region   3089..3157
                /rpt_family="Lx8"
repeat_region   4271..4671
                /rpt_family="CT-rich"
repeat_region   6865..7009
                /rpt_family="L1_MM"
repeat_region   complement(7611..7888)
                /rpt_family="MER63"
repeat_region   9567..9652
                /rpt_family="B4"
repeat_region   complement(10746..11137)
                /rpt_family="MIR3"
repeat_region   complement(11247..11530)
                /rpt_family="MTC"
repeat_region   complement(11715..11855)
                /rpt_family="B4"
repeat_region   11989..12023
                /rpt_family="MIR"
repeat_region   12725..12860
                /rpt_family="(CAAA)n"
repeat_region   12886..12975
                /rpt_family="(TATATG)n"
repeat_region   12985..13215
                /rpt_family="(TATATG)n"
repeat_region   13596..13682
                /rpt_family="(TATG)n"
repeat_region   14133..14174
                /rpt_family="(GGA)n"
repeat_region   14193..14307
                /rpt_family="(TCTCTG)n"
repeat_region   14883..14932
                /rpt_family="(CA)n"
repeat_region   complement(15110..15180)
                /rpt_family="(CA)n"
repeat_region   15838..15911
                /rpt_family="MER33"
repeat_region   17060..17118
                /rpt_family="(GA)n"
repeat_region   17830..18046
                /rpt_family="(TC)n"
repeat_region   18456..18566
                /rpt_family="L1MA5"
repeat_region   18577..18768
                /rpt_family="L1MA5"
repeat_region   complement(19316..19675)
                /rpt_family="B3A"
repeat_region   19677..19709
                /rpt_family="MTD"
repeat_region   19971..20330
                /rpt_family="(CA)n"
repeat_region   20724..20781
                /rpt_family="MTE"
repeat_region   complement(21072..21137)
                /rpt_family="MIR"
repeat_region   21885..22020
                /rpt_family="MIR"
repeat_region   22021..22074
                /rpt_family="B1_MM"
repeat_region   22451..22538
                /rpt_family="(TAAAA)n"
repeat_region   22971..23185
                /rpt_family="Zaphod"
repeat_region   23186..23229
                /rpt_family="URR1B"
repeat_region   24258..24490
                /rpt_family="CYRA11_MM"
repeat_region   26840..26974
                /rpt_family="(CCCCAG)n"
repeat_region   27014..27053
                /rpt_family="CT-rich"
repeat_region   27053..27095
                /rpt_family="(TC)n"
repeat_region   29312..29435
                /rpt_family="(CA)n"
```

```
Query Match      58.1%;  Score 589.6;  DB 10;  Length 162502;
Best Local Similarity  92.7%;  Pred. No. 8.1e-69;
Matches 619;  Conservative  0;  Mismatches  49;  Indels  0;  Gaps  0;

Qy        1 ATGAACTACGTGGGGACACGGGGCGCGGGCGGGCTTACTCGGTGGTGGGGACGGATGGGGGTCTCGCTG     60
            ||||||||||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||||
Db    26384 ATGAACTACGTGGGGACACGGGGCGGGGCGGGCTTACTCGGTGGTGGGGACGGATGGGGGTCTCGCTG  26325

Qy       61 CTGTCCAAACGGGCGCTGGTGGTGCTGCTCTGCCTGTCTGCACAGCGGCGGGACAATCCGCCCAGGCG    120
            ||||||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    26324 CTGTCCAAACGGGCGCTGGTGGTGCTGCTCTGCCTGTCTGCACAGCGGCGGGACAATCCGCCCAGGCG  26265

Qy      121 CCGGGCGCTCTTCACCCTGAACCTCACGTGCGCGGGAACCTGCTTGTCACCGTGGTCAACATG    180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    26264 CCGGGCGCTCTTCACTCTCAACCTCACGTGTGCGCAACCTGCTGTGTGGTTACCGTGGTCAACATG  26205

Qy      241 CTGGCTGCCTTCCTCGACACACCTTTCCTGGCGTGCCCAACTCCACGGTCAGCATGGCGCGCGCTC    300
            |||| ||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    26144 CTGGCCGCCGCCTTCCTCGACACACCTTTCCTGGCCGCCCAACTCCATCGTCAGCGCTGGCGCGCGCTC  26085

Qy      301 AGCATCGACCCGCTGGCGTGGCCGTGTCTTCCCGCGTGAGCGTACCGGCCCAAGATGCGCGCTC    360
            |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    26084 AGCATCGACCCGCTGGCGTGGCTTTCCGCGTGGTGCTTTCCGCGTGAGCGTACCGGTGCCCAAGATGCGCGCTC  26025

Qy      361 CGCGACGCGGCGGCCTCAGTGGTGGCCTACACGTGGCGACGGGCGCTCACCTTCCCAGCCGCC    420
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    26024 CGAGATGCCGCCTTCATTGGTGGCCTACACGTGGCGACGGGCGCTCACCTTCCCAGCCGCC  25965

Qy      421 GCGCTCCGCCCTGTCCTGGCGTCGGCGTTCCACCAGCTGTACGCGCCCCTGTCCAGC    480
            ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    25964 GCGCTCGCCCCTGTCCTGGCTCGGCGTTCCACCAGCTGTATATGCCCTCGTTGCACACCTGTGGCGAGC  25905

Qy      481 CGGGCGGCCAGACGAGCGGCCTGCGCTTCACACTGGCGCCCTTCCACGGCTTCCAGC    540
            |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    25904 CGGGCGGCCCGACGAGCGGCCGCGCTTTGCCGTGTCTTCACCAGCGGCTTCCATGGCGGCTTCCAGC  25845

Qy      541 TTCCTGCTCTCCTTCCGTTCGTTCGTTCGTCGCACGTACCTCAAGGTGGTCGCCGC    600
            |||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    25844 TTCCTCGCTCTCCTTCATCGTTGCTCGTGCTTCACGGTACCTCAAGGTGGCTGGCCCGC  25785

Qy      601 TTCCAATTGCAAGCGCGCATCGACGTGATCACCATGGCGTTGTGTCTTGCTGTTCGGTGGACATA    660
            |||| ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    25784 TTCCACTGCAAGCGCATCGACGTGATCACCACAGCGAGAGCGCTTGTGTGCTTGTTCGGTGGACATA  25725
```

```
Qy          661 CACCCCAG 668
                |||||||||
Db        25724 CACCCCAG 25717
```

```
RESULT 11
LOCUS       AC099826              68539 bp    DNA     linear   HTG 30-JUL-2002
DEFINITION  Papio hamadryas clone RP41-76M19, LOW-PASS SEQUENCE SAMPLING.
ACCESSION   AC099826
VERSION     AC099826.2  GI:22004220
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Papio hamadryas (hamadryas baboon)
  ORGANISM  Papio hamadryas
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Papio.
REFERENCE   1  (bases 1 to 68539)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Papio hamadryas, clone RP41-76M19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 68539)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,R.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigillo,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 68539)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 30, 2002 this sequence version replaced gi:17047165.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
```

```
            --------------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            --------------------- Project Information
                Center project name: Li2534
                Center clone name: 76_M_19
            ---------------------
            * NOTE: This record contains 85 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *
            *      1     684: contig of 684 bp in length
            *    685     784: gap of 100 bp
            *    785    1500: contig of 716 bp in length
            *   1501    1600: gap of 100 bp
            *   1601    2324: contig of 724 bp in length
            *   2325    2424: gap of 100 bp
            *   2425    3092: contig of 668 bp in length
            *   3093    3192: gap of 100 bp
            *   3193    3882: contig of 690 bp in length
            *   3883    3982: gap of 100 bp
            *   3983    4682: contig of 700 bp in length
            *   4683    4782: gap of 100 bp
            *   4783    5480: contig of 698 bp in length
            *   5481    5580: gap of 100 bp
            *   5581    6295: contig of 715 bp in length
            *   6296    6395: gap of 100 bp
            *   6396    7097: contig of 702 bp in length
            *   7098    7197: gap of 100 bp
            *   7198    7916: contig of 719 bp in length
            *   7917    8016: gap of 100 bp
            *   8017    8740: contig of 724 bp in length
            *   8741    8840: gap of 100 bp
            *   8841    9531: contig of 691 bp in length
            *   9532    9631: gap of 100 bp
            *   9631   10356: contig of 725 bp in length
            *  10356   10456: gap of 100 bp
            *  10457   11154: contig of 698 bp in length
            *  11155   11254: gap of 100 bp
            *  11255   11953: contig of 699 bp in length
            *  11954   12053: gap of 100 bp
            *  12054   12736: contig of 683 bp in length
            *  12737   12836: gap of 100 bp
            *  12837   13552: contig of 716 bp in length
            *  13553   13652: gap of 100 bp
            *  13653   14369: contig of 717 bp in length
            *  14370   14469: gap of 100 bp
            *  14470   15190: contig of 721 bp in length
            *  15191   15290: gap of 100 bp
            *  15291   16010: contig of 720 bp in length
            *  16011   16110: gap of 100 bp
            *  16111   16820: contig of 710 bp in length
            *  16821   16920: gap of 100 bp
            *  16921   17587: contig of 667 bp in length
            *  17588   17687: gap of 100 bp
            *  17688   18409: contig of 722 bp in length
            *  18410   18509: gap of 100 bp
            *  18510   19226: contig of 717 bp in length
            *  19227   19326: gap of 100 bp
            *  19327   20044: contig of 718 bp in length
            *  20045   20144: gap of 100 bp
            *  20145   20837: contig of 693 bp in length
            *  20838   20937: gap of 100 bp
            *  20938   21643: contig of 706 bp in length
```

Query Match   57.1%; Score 579; DB 2; Length 68539;
Best Local Similarity 57.9%; Pred. No. 2.5e-67;
Matches 618; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

```
Qy        1 ATGAACTCGTGGGACGGCGGCCCTGGCGGGCGCTACTGGTGGGCACGTGGGCGTCTCGCTG     60
            |||||||||||| || ||||||| || |||| || ||||||||||| |||| ||||| ||||
Db    11322 ATGAACTCGTGGGACAGCGGCCGCGGCCTGCGTGGTGGGGCACCATGGGCGTTCGCTG    11381

Qy       61 CTGTCCAACGCGCGCGGTGGTGCTGCTCGCTGCCTGCACAGCGGCGCATCCGCCCGCCCAGGCG    120
            ||||||||||||||| ||||||||||||| |||||||||||||||||||||||||||||||||||
Db    11382 CTGTCCAACGCGCGGCTGGTGCTGCTGCCTGCGCACAGCGGCGCATCCGCCCGCCCAGGCG    11441
```

RESULT 12
LOCUS   AY255585   525 bp   mRNA   linear   ROD 08-APR-2003
DEFINITION Mus musculus G protein-coupled receptor GPR26 mRNA, partial cds.
ACCESSION   AY255585
VERSION   AY255585.1 GI:29611571
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 525)
  AUTHORS   Vassilatis,D.K., Hohmann,J.G., Zeng,H., Li,F., Ranchalis,J.E.,
            Mortrud,M.T., Brown,A., Rodriguez,S.S., Weller,J.R., Wright,A.C.,
            Bergmann,J.E. and Gaitanaris,G.A.
  TITLE   The G protein-Coupled Receptor Repertoires of Human and Mouse
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (8), 4903-4908 (2003)
REFERENCE   2 (bases 1 to 525)
  AUTHORS   Vassilatis,D.K., Hohmann,J.G., Zeng,H., Li,F., Ranchalis,J.E.,
            Mortrud,M.T., Brown,A., Rodriguez,S.S., Weller,J.R., Wright,A.C.,
            Bergmann,J.E. and Gaitanaris,G.A.
  TITLE   Direct Submission
  JOURNAL   Submitted (14-MAR-2003) Primal, Inc., 1124 Columbia Street,
            Seattle, WA 98104, USA
FEATURES   Location/Qualifiers
     source   1..525
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     CDS   <1..>525
                     /codon_start=1
                     /product="G protein-coupled receptor GPR26"
                     /protein_id="AAO85097.1"
                     /db_xref="GI:29611572"
                     /translation="KMRLRDAFMVAYTWLHALTFPATALALSWIGFHQLYASRTLCS

```
Qy      121 CCGGCGCGTTCTCACCCTGAACGTGCGCGGAACCTGCTGTGCAACTGGTGTCACGTGTCAACATG    180
            |||||||||| || ||||||||| |||||||||||||||| | |||||||||||||||||||||||
Db    11442 CCGGCGCGTTCTTCACCCTGAACCTGCGGGAACCTGCTGCGTGTGCAACTGGTGTCAACATG    11501

Qy      181 CCGCTCACGCTGTGGCCCGGCCCGGTCGTGTGGGCCCAGCGGCGGGGGGGGGACCGGCTGTGCCGC    240
            |||||||||||||| |||||||||||||| |||||||||||||||||||||||||||||||| ||||
Db    11502 CCGCTCACGCTGGCGGGCCGGGCCGGTCGTGTGGGCCAGCGGCCAGCGGACCGGCTGTGCCGC    11561

Qy      241 CTGGCCTGCCTTCCTCGACACCTTTCCGTGGCTGCGCCAACTCCATGTCCAGCATGCCGCGCCTC    300
            |||||||||||||||||||| ||||||||||||| ||| |||||||||||| |||||||||||||
Db    11562 CTGGCCTGCCTTCCTCGACACCTTTCCGTGGCTGCGCCAACTCCATGTCCAGCATGCCGCACTC    11621

Qy      301 AGCATCGACCCGCTGGGGTTGGGGTTGGGGTGTCTTCCGGCTGAGCTACCGGCGCAAGATGCCGCTC    360
            |||||| || ||||||||| ||||| |||| ||| ||||||||||||| ||||| |||| ||||||||
Db    11622 AGCATCGACCGCTGGGGGTTGGGGCCGTGGTCTTCCGGCTGAGCTACCGGCGCAAGATGCCGCTC    11681

Qy      361 CCGCGACGGCGGCGGCCTCATGGGTGGCCTACACGTGGGCGCAGCGGCGCTCACCTTCCCAGGCGCCC    420
            |||||||||||||||||||| ||||| ||| |||||| |||| ||||||||||||||| ||||| ||||
Db    11682 CCGCGACGGCGGCGGCCTCATGGGTGGCCTACACGTGGGCGCAGCGGCGCTCACCTTCCCAGGCGCCC    11741

Qy      421 GCGCTCGCCCCTGTTCCTCGGGCGTTCCCACCAGCGTGTACGGCGCTCGTTCCAGCGTTGTGCAGC    480
            |||| | ||||| |||| |||| |||||| || ||||| |||| ||||||||||| ||||| |||||
Db    11742 GCGCTCGCCCCTGTTCCTCGGGCGTTCCCACCAGCGTGTACGGCGCTCGTTCCAGCGTTGTGCAGC    11801

Qy      481 CGGCGGCCAGACGAGCGGCGGCCTGCGCTTCGGCGCCGTTCCTCACTGTGGGGCGCCTTCCAGCGTGCTCAGC    540
            |||| ||| || |||||||| |||||||||| ||||||||||||| ||||||||||||| ||||| |||||
Db    11802 CGGCGGCCGGACGAGCGGCGGCCTGCGCTTCGGCGCCGTTCCTCACTGTGGGGCGCCTTCCAGCGTGCTCAGC    11861

Qy      541 TTCCT-GCTCTCCTTCGGTCGTCGTGTGCTCTGCGCACGTACCTCAA-GGTGCTCAAGGGTGGCCC    598
            |||||  |||||||||| |||||| |||||||| ||| |||||| |||  ||||| |||||||||||
Db    11862 TTCCTGGCTCTCCTTCGGTCGTCGTGTGCTCTGCGCACGTACCTCAAGGTGCTCAAGGGTGGCCC    11921

Qy      599 GCTTCCATTGCAA-GGGCATCGACGTGATCA    628
            |||||||||||||  |||||||||| |||||
Db    11922 GCTTCCATTGCAAGGGCATCGACGTGATCA    11952
```

Contig/gap structure (Db positions):

```
*   21644   21743:  contig of 697 bp in length
*   21744   22440:  gap of 100 bp
*   22441   22540:  contig of 100 bp in length
*   22541   23254:  gap of 714 bp
*   23255   23354:  contig of 100 bp in length
*   23355   24029:  gap of 675 bp
*   24030   24129:  contig of 100 bp in length
*   24130   24852:  gap of 723 bp
*   24853   24952:  contig of 100 bp in length
*   24953   25672:  gap of 720 bp
*   25673   25772:  contig of 100 bp in length
*   25773   26488:  gap of 716 bp
*   26489   26588:  contig of 100 bp in length
*   26589   27282:  gap of 694 bp
*   27283   27382:  contig of 100 bp in length
*   27383   28101:  gap of 719 bp
*   28102   28201:  contig of 100 bp in length
*   28202   28917:  gap of 716 bp
*   28918   29017:  contig of 100 bp in length
*   29018   29726:  gap of 709 bp
*   29727   29826:  contig of 100 bp in length
*   29827   30513:  gap of 687 bp
*   30514   30613:  contig of 100 bp in length
*   30614   31314:  gap of 701 bp
*   31315   31414:  contig of 100 bp in length
*   31415   32128:  gap of 714 bp
*   32129   32228:  contig of 100 bp in length
*   32229   32949:  gap of 721 bp
*   32950   33049:  contig of 100 bp in length
*   33050   33778:  gap of 729 bp
*   33779   33878:  contig of 100 bp in length
*   33879   34606:  gap of 728 bp
*   34607   34706:  contig of 100 bp in length
*   34707   35425:  gap of 719 bp
*   35426   35525:  contig of 100 bp in length
*   35526   36242:  gap of 717 bp
*   36243   36342:  contig of 100 bp in length
*   36343   37035:  gap of 693 bp
*   37036   37135:  contig of 100 bp in length
*   37136   37833:  gap of 698 bp
*   37834   37933:  contig of 100 bp in length
*   37934   38626:  gap of 693 bp
*   38627   38726:  contig of 100 bp in length
*   38727   39431:  gap of 705 bp
*   39432   39531:  contig of 100 bp in length
*   39532   40246:  gap of 715 bp
*   40247   40346:  contig of 100 bp in length
*   40347   41058:  gap of 712 bp
*   41059   41158:  contig of 100 bp in length
*   41159   41881:  gap of 723 bp
*   41882   41981:  contig of 100 bp in length
*   41982   42705:  gap of 724 bp
*   42706   42805:  contig of 100 bp in length
*   42806   43504:  gap of 699 bp
*   43505   43604:  contig of 100 bp in length
*   43605   44327:  gap of 723 bp
*   44328   44427:  contig of 100 bp in length
*   44428   45137:  gap of 710 bp
*   45138   45237:  contig of 100 bp in length
*   45238   45933:  gap of 696 bp
*   45934   46033:  contig of 100 bp in length
```

```
RRPDERLRFAVFTSAFHALSFLLSFIVLCFTYLKVARFHCKRIDVITMQTLVLLV
DIHPSVRERCLEEQKRRQRATKKISTFIGTFLVCFAPYVITRLIVELFSTAPIGSHWG
VLSKCLIAYSKAASDP"

ORIGIN

Query Match          45.0%;   Score 456.2;  DB 10;  Length 525;
Best Local Similarity  91.8%;  Pred. No. 1.2e-50;
Matches  482;  Conservative  0;  Mismatches  43;  Indels  0;  Gaps  0;

Qy   349 AAGATGGCGCGCCCGGACCCGACCCGCCGCGGCCGCTGCTGGCTGCACGCGGCCGCTCACC 408
         ||||||||| ||||| |||||||||||| |||||||||||||||||||||||||||||||
Db     1 AAGATGGCGGCCTCCGAGATGCGCCGCTTCATCGGCTACTGGCCTGCACGCGGCCGCTCACC 60

Qy   409 TTCCCAGCCGCGCCGCGCTCGCCCTCTTCGGCTACGGTGTACGGCTCGTGC 468
         |||||||||||||| ||| |||||| |||||||| |||||||||||||
Db    61 TTCCCGGCCACCGGCGCTCCGCCTTCCAACAGCTATATGCCTTCGCGC 120

Qy   469 ACGGTGTGCGAGCGCGGCCCGGAGCCGCCGCTTCGGCCGTCTTCACTGGCCCTTC 528
         |||||||||||||||||||||||||||||||||||||||||||||
Db   121 ACACTGTGCAGCCGCGGGCCGGACCAGCCGCTTTGCTGTCTTCACCAGCCCTTC 180

Qy   529 CACGGCTCTCAGCTTCCTGCTCTCTTCGTCGTGTCGTCGCAGGTACCTCAAGGTGCTC 588
         |||||||||||||||| ||| ||||||||| |||| |||||||||||||||
Db   181 CATCGCCGCTCAGCTTTCCGTCTCTCCTTCATCGTGCTCGTTGTGTCGCAGGTGCTC 240

Qy   589 AAGGTGGCGCCCGGCCTTCCATTGCAACGGCATCGACGTATCGAGCAGCCTGGTGCTG 648
         |||| |||||||||||||| |||||||||| ||||||||||||||||||||||||
Db   241 AAGGTGGCCCCGCGTTCCCACTCACAGCATCAACCATGGACGACGCTGTGTGCTG 300

Qy   649 CTGGTGGACCTGCACCACCGTGCACCCCAGTGTCGTCTCTGGAGAGAGAAGCGGGAGGCGA 708
         ||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   301 TTGGTGGACATACACCCAGTGTGAGGGAACGGTCTCTGGAGGAACAGAAGCGGCAGGCGA 360

Qy   709 CAGGCGAGCCACCCAAGAAAGAATCAGCACAACCTTCCTGGTTCGTGTGTGTGCTTCGCGCC 768
         |||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Db   361 CAGGCTGCCACCAAGAAGATCAGCACCTTCATCGGTACCTTCCTGTGTTTGCCACCC 420

Qy   769 TATGTGATCACCCAGGCTAGTGAGGAGCTCTTCTCCAGGTCGCCATCGGCTCCCACTGGGGG 828
         |||||||||| || ||||||||||||||||||||||||||||
Db   421 TATGTGATTACCAGGCCGTGGGTGGAACTCTTCTCCACAGCACCCATTGGCTCTCACTGGGGA 480

Qy   829 GTGCTGTCCAAGTGCTTGGCGTACAGCAAGGCCGCCCCATCCGACCCC 873
         |||||||||||||| |||||||||||||||||| |||||||||
Db   481 GTGCTGTCCAAGTGCTTGGGCCTACAGCAAGGCCGCGTTCTGACCCC 525

RESULT 13
LOCUS       AX148190                1092 bp    DNA     linear   PAT 08-JUN-2001
DEFINITION  Sequence 31 from Patent WO0136471.
ACCESSION   AX148190
VERSION     AX148190.1  GI:14347088
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chen,R., Dang,H.T. and Lowitz,K.P.
  TITLE     Endogenous and non-endogenous versions of human g protein-coupled
            receptors
  JOURNAL   Patent: WO 0136471-A 31 25-MAY-2001;
            Arena Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..1092
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match          41.2%;   Score 418.2;  DB 6;  Length 1092;
Best Local Similarity  64.5%;  Pred. No. 9.3e-46;
Matches  624;  Conservative  0;  Mismatches  343;  Indels  0;  Gaps  0;

Qy     1 ATGAACTCGTTGGGACGGCGGGCGCCTGGCGGGGGGCTACTGGTGGGGACACCGATGGGCGGTCTTCGCTG 60
         |||| ||| |||||||||||||||| ||||||||||||||||||||||||||||||
Db     1 ATGGGCCCCCGGCGGAGGCGGCCGCTCCTGGCGGGGTTCCTCGGGTGATTGGGGTACTGGGGCGGTG 60

Qy    61 CCGGGGGCTCTTCACCCTGAACCTCAGCTGGGGGGAACCTGCGGCCACAGCAGGCGCAGTGGCTCAACATG 180
         ||| ||| ||||||||||||||||||||||| ||||||||||||||||||
Db    61 CTATCCAACGCACTGGTGGCTGTTGTGTTTGTTGGGCTTACAGGCCTGCGGTTCCCGCACTGGAGCC 120

Qy   121 CCGGCTCACGCTGCGCCGGCCGGTCGTGCGCAGCGGCAGCCACCTGCTGCTTCGGCGGCCGCTGGTCGCC 180
         |||||||| |||| ||||||||||||||||||||| |||||||
Db   121 TCAGGCGGTCCTCCTGGTGGAATCTGTGGCTGCTGGCCGACATG 180

Qy   181 CCGCTCACGCTGCGCCGGCCGGTCGTGCGCAGCGGCAGCCACCTGCTGCTTCGGCGGCCGCTGGTCGCC 240
         ||| ||||||||||||||||||||||||||| ||||||||
Db   181 CCCTTCACGCTGCTCCGGTCGGTTGCAGTGGGCGGGGGGACAGCCGGGGGCATGCCAA 240

Qy   241 CTGGCTGCCTTTCCTCGACACCTTCCTGGCTGCCAACTCCATATGCTTCAGCATGGCCGCGCTC 300
         ||||| ||||||||| |||||||||| |||| |||||||||||||
Db   241 GTCATTGGCTTCCTGGACACCTTCCTGGGCGTCCAAGCGGGGTGGGGGCTGGTG 300

Qy   301 AGCATGGACCGCTGGGGTGGCCGGTGCTCTTCCCCGGCTACACCGGGGCCCAAGATGTGCCCTC 360
         |||||||||||||||| |||||||||||||||||||||||||||||||
Db   301 AGCCGAGACCAAGTGGCATTGGCCGGTGGGGGTCCCACGCCTTACCGGGCCCCCTGGACCG 360

Qy   361 CGGCGACGGCGCGCTGCAGGTGGCCTACACGTGGGTGGCCTTCCCACTGGCCGCTCACTGGCTGTGCTG 420
         |||||||| |||||||||||||||||||||||||||||||||||
Db   361 CGCTARGCCGGCACGCCCGTGCTGCTGGGGGGACAGTGCGCTGGCCTTTCAGGGCT 420

Qy   421 GCGCTCGCCCCTGTCCCGTGGGTCTTCCCACCAGTGTGTACCCCTGTCGCAGTGTGCAGC 480
         |||||||||||||||||||||||||||||||||| |||||||
Db   421 GCACTTGGCTGCTCGTGGCTTGGGTCGACGGGTTCCGTTGGCACGCGGTTGGCCACGC 480

Qy   481 CGGGCGGCCAGACGAGGCGCCTGCGCTTCGCGCCTCTTCACTGGCGGCTCTCAGC 540
         |||||| ||||||||| ||||||||||||||||||
Db   481 CCGCCCGAGCCTGACGTCGGGCGTCCGGCGCCCCAGCCTTCACCGGGGGC 540

Qy   541 TTCCTGCTCTCCTTCGTCGTGCTGCTGCTGCACGTACCTCAAGGTGCTCAAGGTGGCCGCGC 600
         ||||| ||||||| |||||||||||| ||||||||||||
Db   541 TTCGTGCCTGCCGCGGTGCTGTGCCCGGTTGCTGCACCTCCAGGTGGCCAGGC 600

Qy   601 AGCCACTGCCAGCGCCATGGACACCGTCACCATGAAGGCGTCGCGCTGCTGCTGGTGGACCTG 660
         ||||||||||||||||| ||||||||||||||||||||||||||||
Db   601 GCCACTGGCTGTCGTGGCCTGCGTTTGGCTGCTGCGCTCCTGCGCCTG 660

Qy   661 CACCCCAGTGTGCGGGGAACGGTGTCTGGAGGAGCAGAAGCGGGAGGCGACGGAGGCACC 720
         ||||||||||||||||||||||| ||||||||||||||
Db   661 CACCCCAGTGTGCGGGAGCGCTGCCTCGAAGAGCAGAAGCGGCGGCAGCGCGCCACCAAG 720

Qy   721 AAGAAGAATCAGCACCTTCATAGGGACTTCCTTGTGTGCTTCGCGCCCTATGTGATCACC 780
         |||||||| ||||| ||||| |||||||||||||||||||||
Db   721 AGGAAGAATTGGCATTGCTATTGCGACCTTCATCTGCTTTGCCCCGTATGTCATAGACC 780

Qy   781 AGGCTAGTGGAGGCTCTTCTCCACGCGTGCCACTGGGGGGTGCTGTCGTCCAAG 840
         |||||||||||||||| ||||||| |||||||||||||||||||
Db   781 AGGCTGGCGGAGCTCGTGCCCTTCGTCAACGGCATCCTCAGCAAG 840

Qy   841 TGCTTGGCGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGGCACCAGG 900
         ||||||||| || ||||| ||||||| |||||| |||||||||
Db   841 TGCCTGACCTACAGCAAGGCGGTGGCCGGTACTCTGCTCCGCGCCG 900

Qy   901 TACCGGCAAAAGCTGCAAAGGAGATTCTGAACAGGCTCCTGCACAGACGCTCCATCCACTCC 960
         |||||| |||||| || ||||||||||||||||||
Db   901 TTCCGGCCAAGTCCCTGGCGGGCCATGGTGCACCGGCGCTCGCTGAAGAGAACCCCGCCAGCA 960

Qy   961 TCTCTGGCC 967
         |||||
Db   961 TCCACCC 967

RESULT 14
LOCUS       AX592452                1092 bp    DNA     linear   PAT 27-JAN-2003
```

DEFINITION   Sequence 9 from Patent WO0232962.
ACCESSION   AX592452
VERSION   AX592452.1   GI:27950554
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Glucksmann,M.A. and Meyers,R.
  TITLE   80090, 52874, 52880, 63497, and 33425 methods and compositions of
             human proteins and uses thereof
  JOURNAL   Patent: WO 0232962-A 9 25-APR-2002;
             Millennium Pharmaceuticals, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..1092
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Query Match   41.1%;   Score 416.6;   DB 6;   Length 1092;
Best Local Similarity   64.4%;   Pred. No. 1.5e-45;
Conservative 623;   Mismatches 344;   Indels 0;   Gaps 0;

Qy   1 ATGAACTCGTGGGGACGCGCGGC...   60
Db   1 ATGGGCCCCGGACGGAGCGC...   60

Qy   61 CTGTCCAAGCGGCGTCGTTG...   120
Db   61 CTATCCAACCGCACTCGTGC...   120

Qy   121 CCGGGCGCTCTTCACCCTGA...   180
Db   121 TCAGGCGTCTCCTCCTGGTG...   180

Qy   181 CCGCTCACGCTGGCCGGCCG...   240
Db   181 CCCTTCACGCTGCTCGGTGA...   240

Qy   241 CTGGCGCTGCCTTCCTCCGAC...   300
Db   241 GTCATTGGCTTCCTCTGGAC...   300

Qy   301 AGCATCGACCGCTGGGTGGGC...   360
Db   301 AGCGGCGAGACCAGTGGCGC...   360

Qy   361 CGGCGACGCGCGGCGCTCATG...   420
Db   361 CGCTATGCCGCCGGCCTCTGC...   420

Qy   421 GCGCTCGCGGCCCTGTTCCTG...   480
Db   421 GCACTTGGCTGCTCGTGGCTT...   480

Qy   481 CGGGGCGCCAGACGAGCGGCC...   540
Db   481 CCGGCCCGAGCCTGAGGCGGT...   540

Qy   541 TTCCTGCTCTCCTTCGGTCGTG...   600
Db   541 TTCGTGCTGCTGCCGCGGTCC...   600

Qy   601 TTCCATTGCGAAGCGCCATGC...   660
Db   601 AGCCACTGCCAGCACTCTCAC...   660

Qy   661 CACCCCAGTGTGGCGGGAAGC...   720
Db   661 CACCCCAGTGTGGCGGGCAGG...   720

Qy   721 AAGAAGAATCAGCACTTCATAGGAGC...   780
Db   721 AGGAAGAATTGGCGATTGCTATTGCG...   780

Qy   781 AGGCTACTGATGGAGCTCTTCTCCA...   840
Db   781 AGGCTGGCGGAGCTCGTGCCCTTC...   840

Qy   841 TGCTTGGCGTACAGCAGCAAGGCCG...   900
Db   841 TGCCTGACCTACAGCAAGGCCGTGG...   900

Qy   901 TACCGCGAAAAGCTGCGAAGGAGAT...   960
Db   901 TTCCGGCCAAGTCCTGGCGCGGCAT...   960

Qy   961 TCTGGGCC   967
Db   961 TCCACCC   967

RESULT 15

LOCUS   AX592450   1352 bp   DNA   linear   PAT 27-JAN-2003
DEFINITION   Sequence 7 from Patent WO0232962.
ACCESSION   AX592450
VERSION   AX592450.1   GI:27950552
KEYWORDS
SOURCE   Homo sapiens (human)
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Glucksmann,M.A. and Meyers,R.
  TITLE   80090, 52874, 52880, 63497, and 33425 methods and compositions of
             human proteins and uses thereof
  JOURNAL   Patent: WO 0232962-A 7 25-APR-2002;
             Millennium Pharmaceuticals, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..1352
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS   210..1301
                     /codon_start=1
                     /note="unnamed protein product"
                     /protein_id="CAD62019.1"
                     /db_xref="GI:27950553"
                     /translation="MGPEALLAGLLVWVLAVALLSNALVLLCCAYSAELRTRASGVL
                     LVNLSLGHLLLAALDMPFTLLGVWNRGRTPSAPGCAQVIGFLDTFLASNAALSVAALSA
                     DQWLAVGFPLRYAGRLRPRYAGLLLGCAWGQSLAFSGAAIGCSWLGYSSAFASCSLRL
                     PEPEPRPRPAAFTATLHAVGFVLPLAVLCLTSLQVHRVARSHCQRMDTVTMKALALLA
                     DLHPSVRQRCLIQQKRRHRATRKIGIA1ATFLICFAPYVMTRLAELVPFVTNNAQWG
                     ILSKCLTYSKAVADPFTYSLLRRPRQVLAGMVHRLLKRTPRPASTHDSSLDVAGMVH
                     QLLKRTPRPASTHNGSVDTENDSCLQQTH"
ORIGIN

Query Match   41.1%;   Score 416.6;   DB 6;   Length 1352;
Best Local Similarity   64.4%;   Pred. No. 1.4e-45;
Conservative 623;   Mismatches 344;   Indels 0;   Gaps 0;

Qy   1 ATGAACTCGTGGGGACGCGCGGC...   60
Db   210 ATGGGCCCCGGACGGAGCGC...   269

Qy   61 CTGTCCAAGCGGCGTCGTTG...   120
Db   270 CTATCCAACCGCACTCGTGC...   329

Qy   121 CCGGGCGCTCTTCACCCTGA...   180
Db   330 TCAGGCGTCTCCTCCTGGTG...   389

```
Qy   181  CCGCTCACCGCTGGCGCGGCCGTCGTGCGCAGCGGCGGCCGGGGGCGACCGCCCTGTGCCGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   390  CCCTTCACCGCTGCTCGGCGTGTGATCGGCGGCGGCGGAGACACCGTCGGGCGCCCGGCATGCCAA 449

Qy   241  CTGGCTGCCTTCCTCCGACACCTTCCTGGCTGCCAAACTCCATGCTCAGCATGGGCGGGCTC 300
          ||||||  |||||||||||||  |||||||
Db   450  GTCATTGGCTTCCTTGGGACACCTTCCTGGCGGGTCCAAACGGCGGCTGAGCCGTGGGCGGCTG 509

Qy   301  AGCATCGACCGCTGGGGTGCCGGTGGTCTTCCCGCTGAGCTACCGGGCCAAAGATGGCGCCTC 360
          ||||
Db   510  AGCCCAGACCAGTGGCTGGCAGTGGGCTTCCCACTGCGGCTACCGCCGGACGGCCTCCGACCG 569

Qy   361  CGGGACGCGGGCGGCCTCATGGTGGGCTACAGTGGGCTGCAGCGCTCACCTTCCCAGCCGCC 420
          ||||||        ||||||||||||||||||||
Db   570  CGCTTATGCCGGCCGGCCTGCTGCGTGCCTGGGGACAGTGGCTGGCCTTTCTCAGGCGCT 629

Qy   421  GCGGCTGCGCCCTGTCCTGGCGTCGCGCTTCACCAGCTGTACGCCTGGTGCACGCTGTGCAGC 480
          ||||||  |||||||||||        ||||||||||||
Db   630  GCACTTGGGCTGCTGCTGGGCTTTGGCTACACGGAGCGGCCTTCGTCGTGCTCCGGTTCGGCTCGCCTG 689

Qy   481  CGGGCGGCCAGACGAGCGCCTGCGCGCTTCGCGCGTCTTCACTGGGCCGCCTTCCAGCTCTCAGC 540
          ||||||      |||        |||||||||||||||
Db   690  CCGCCCGAGCCTGAGCGTCCGCGCTTGCGGCTGTACCCCACGCCTCCACTGGCGTGCAGGCTG 749

Qy   541  TTCCTGCTCTCCTTCGTCGTGCTCCTGCTCGCACGTACCTCAAGGTTGCTCAAGGTGGCCCGC 600
          ||||  |||||||||||||||||||||||||||||||||||
Db   750  TTCGTGCTGCCCGGCTTGGCGGTGCTTCTGCTCGTTCACCTTCGCTTCCAGGTGGCAGCC 809

Qy   601  TTCCATTGCAAAGCGCCATCGACGTGATCACCATGCAGACCGCTGGTGTCGTGGGTGGACCTG 660
          |||||  |||||||||||||||        ||||||||||||
Db   810  AGCCACTGCCAGCCGCATGGACACTGTCACCATGAAGGGCGGCTTCGCGCCGGACCTG 869

Qy   661  CACCCCCAGTGTGCGGGAACGCTGTCTGGAGGAGCAAGAGCGGGGAGGCGGACGAGCGCCACC 720
          |||||||||||||||||||||||
Db   870  CACCCCCAGTGTGCCGCAGCCGCCTGCTCATCCAGCAGAAGCGGCGCGCGCCCACCGCCACC 929

Qy   721  AAGAAGATCAGCACCTTCATPAGGGACCTTCCTTGTGTGGCTTCGGCCCCTAGTGATTCACC 780
          ||||||||||||  |||||||||||
Db   930  AGGAAGATTGGCATTGCTATTGCCGAACCTTCCTCATCTGCTTTGCCCCGTPATGTCATGACC 989

Qy   781  AGGCTPAGTGGAGCTCTTCTCCACGTGCCCATCGGCTCCATCGGGGGTGCTGTCCAAG 840
          ||||||  |||||||||
Db   990  AGGGTGGGCGGGAGCTCGTGCCCCTTCGTCACCGTGAACGCCCCAGTGGGGCATCCTCAGCGAAG 1049

Qy   841  TGCTTPGGGCGTACAGCAAGAGCGGCATCCGACCCCTTTGTGTACTCCTTACTGCGACACCAG 900
          ||||
Db   1050 TGCCTGACCTPACAGCAAAGGGGGTGGCCGAACCGGTTCACGTPACTCTTCTGCTCCGCGGCCG 1109

Qy   901  TACCCCAAAAGCTGCAAAGGAGAGATTCTGAACAGGCTCCTGCACAGAGGCTTCCATCCACTCC 960
          ||||||  |||
Db   1110 TTCCCGCCAAGTCCTCGGCCGGCCATGGTGCACCGGGCTGCTGAAGAGAACCCCGGCGCCCAGCA 1169

Qy   961  TCTGGCC 967
          ||| ||
Db   1170 TCCACCC 1176
```

QM nucleic - nucleic search, using sw model

```
                    GenCore version 5.1.6
                    Copyright (c) 1993 - 2005 Compugen Ltd.
```

Run on: July 5, 2005, 12:06:10 ; Search time 3906 Seconds
                                            9881.508 Million cell updates/sec

Title: US-10-735-991-5
Perfect score: 1014
Sequence: 1 atgaactcgtgggacgcggg..........ttctgcgggtgtctgagtga 1014
Scoring table: IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 34395544 seqs, 1903213470 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   EST:*

```
                 1: gb_est1:*
                 2: gb_est2:*
                 3: gb_htc:*
                 4: gb_est3:*
                 5: gb_est4:*
                 6: gb_est5:*
                 7: gb_est6:*
                 8: gb_gss1:*
                 9: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	961.4	94.8	963	9	AY420105	AY420105 Homo sapi
2	871.6	86.0	2802	9	AK036100	AK036100 Mus muscu
3	825.4	81.4	963	9	AY420107	AY420107 Mus muscu
4	557.2	55.0	666	9	AY420106	AY420106 Pan trogl
5	538.8	53.1	741	7	CO934052	CO934052 AGENCOURT
6	521.8	51.5	736	7	CO934490	CO934490 AGENCOURT
7	517.8	51.1	730	7	CO934461	CO934461 AGENCOURT
c 8	491.8	48.5	1211	3	AK042755	AK042755 Mus muscu
9	491.8	47.7	861	6	CB528011	CB528011 UI-M-FY0-
10	483.4	47.4	703	8	BB639955	BB639955 BB639955
c 11	480.8	38.7	537	8	BH118705	BH118705 RPCI-24-3
12	392.4	32.1	1005	9	CN503VDR	CN503VDR Tetraodon
13	325.4	29.6	318	7	T09060	T09060 EST06953 In
14	299.8	26.6	1081	9	CN503YUE	CN503YUE Tetraodon
15	270	24.1	870	6	CN502SGY	CN502SGY Tetraodon
16	244.4	23.8	739	6	CB521501	CB521501 UI-M-GH0-
17	241.6	23.1	601	9	CN501YOS	CN501YOS Tetraodon
18	234.6	18.2	609	9	AL172269	AL172269 Tetraodon
19	184.2	16.9	601	9	AL123985	AL123985 Fugu rubr
c 20	161	15.9	684	9	AL253255	AL253255 Tetraodon
c 21	150.8	14.9	650	9	BY725178	BY725178 BY725178
c 22	141.4	13.9	678	9	DR43G7S	DR43G7S Danio rer
c 23	115.6	11.4	648	2	BB628814	BB628814 BB628814
c 24	112.8	11.1	931	5	BU216839	BU216839 603104803
25	110.4	10.9	747	5	BX165029	BX165029 Danio rer
26	109.2	10.8	855	5	BU369227	BU369227 60597413
27	104.8	10.3	569	6	CB113745	CB113745 tigr-gss-
28	86	8.5	815	9	CB989325	CB989325 AGENCOURT
29	79	7.8	889	7	CN838171	CN838171 AGENCOURT
30	79	7.8	890	7	CN842565	CN842565 AGENCOURT
31	79	7.8	906	7	CN839221	CN839221 AGENCOURT
32	79	7.8	1113	9	AY420480	AY420480 Homo sapi
33	78.4	7.7	1100	9	AY420481	AY420481 Pan trogl
34	76.2	7.5	721	7	CF147813	CF147813 AGENCOURT
35	75.2	7.4	853	9	AY409236	AY409236 Homo sapi
36	73.8	7.3	732	7	CF147822	CF147822 AGENCOURT
37	73.6	7.3	925	9	CN5009IP	CN5009IP Drosophil
38	73	7.2	903	9	CN502QPU	CN502QPU Tetraodon
c 39	71.6	7.1	1548	9	AG435551	AG435551 Mus muscu
c 40	71.4	7.0	908	9	CN503MK4	CN503MK4 Tetraodon
c 41	71	7.0	1038	3	BC017857	BC017857 Homo sapi
42	71	7.0	2020	3	CR601077	CR601077 full-leng
43	70.8	7.0	697	9	CG248403	CG248403 CG2WIH29TH
44	70.2	6.9	803	7	CK146110	CK146110 AGENCOURT
45	69.2	6.8	975	9	CN5039RV	CN5039RV Tetraodon

ALIGNMENTS

RESULT 1
LOCUS   AY420105
DEFINITION   AY420105   963 bp   DNA   linear   GSS   12-DEC-2003
               Homo sapiens HCM7113 gene, VIRTUAL TRANSCRIPT, partial sequence,
               genomic survey sequence.
ACCESSION   AY420105
VERSION   AY420105.1   GI:39776062
KEYWORDS   GSS.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 963)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
               Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
               Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
               Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
               gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
  PUBMED   14671302
REFERENCE   2 (bases 1 to 963)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
               Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
               Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
               Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
               Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering them
               based on alignment.
FEATURES   Location/Qualifiers
     source   1..963
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     gene   <1..>963
                     /locus_tag="HCM7113"
ORIGIN

Query Match   94.8%; Score 961.4; DB 9; Length 963;
Best Local Similarity   99.9%; Pred. No. 1.2e-191;
Matches 962; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   52 GTCTCGGCTGCTGTGTCCAACCGCGCGGCTGGTCTGCCTCTCGCTCGGCGTCGCAGCGCGGCAGCGGGACACATCCGC 111
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GTCTCGGCTGCTGTGTCCAACCGCGCGGCTGGTCTGCCTCTCGCTCGGCGTCGCAGCGCGGCAGCGGGACACATCCGC 60

```
Qy    112 CGCCAGGCGGCGCCGGGCGGCTCTTCACGCTTGACCTTCACGTGCGGGAGACCTGCTGTGCACCGTG 171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CGCCAGGCGGCGCCGGGCGGCTCTTCACGCTTGACCTTCACGTGCGGGAGACCTGCTGTGCACCGTG 120

Qy    172 GTCAACATGCCGCTCACGCTGGGCGGCGGCCGAGCGCCGGGCGCAGGGCGGCGGCGGCGGCGACCGC 231
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTCAACATGCCGCTCACGCTGGGCGGCGGCCGAGCGCCGGGCGCAGGGCGGCGGCGGCGGCGACCGC 180

Qy    232 CTTGTGCCGCCTGGCGTGTGCCTTCCTCGACACACCTTCCATTGCTCATGCTCAGGCATG 291
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTTGTGCCGCCTGGCGTGTGCCTTCCTCGACACACCTTCCATTGCTCATGCTCAGGCATG 240

Qy    292 GCCGCGCTCAGCATCGACCGCTGGGTGGTGTTCTTCCGCGGCTCAGCGTACCGGCCCAAG 351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCGCGGCTCAGCATCGACCGCTGGGTGGTCCGCGTGTTCTTCCCGCGGCTACCGGCCCAAG 300

Qy    352 ATGGCGGCCCTTCCGGACGGGGCGGGCCGCTCATGGCGGCCCTACACGTTGGCTGCAGCATG 411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATGGCGGCCCTTCCGGACGGGGCGGGCCGCTCATGGCGGCCCTACACGTTGGCTCACCTTC 360

Qy    412 CCAGCGCGCGCCGCTCGCCCTGTCTGGGCTCGGGCTTCCACCAGCTGTACGGCCTCGTGACG 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CCAGCGCGCGCCGCTCGCCCTGTCTCGGGCTCGGGCTTCCACCAGCTGTACGGCCTCGTGACG 420

Qy    472 CTGTGCCAGCGCCGGCCGGACGGCGCGGGCCCTGGCTTCGCGCCTTCACTGCGCCTTCCAC 531
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTGTGCCAGCGCCGGCCGGACGAGCGGCGGCGTCTTCGCGCCTTCACTGGCGCCTTCCAC 480

Qy    532 GCTCTCAGCTTCTGCTCTCTCGTTCGTCGTGCTCGCGTACTCAAGGTGCTCAAG 591
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GCTCTCAGCTTCCTTGCTCTCTCGTTCGTCGTGCTCGCGTACTCAAGGTGCTCAAG 540

Qy    592 GTGGGCCGGCTTCCATTGCAAGGCGCGCATCGAGGTGATCACGACACGCTGGGTGCTGCTG 651
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTGGGCCGCTTCCATTGCAAGCGCCATCGAGGTGATCACCATGCAGACGCTGGGTGCTGCTG 600

Qy    652 GTGGGACCTGCACCCCAGTGTGTGCGGGAAGCGCTGTCTTGTGGAGAGGACAGAAGCGGACAG 711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GTGGGACCTGCACCCCAGTGTGTGCGGGAAGCGCTGTCTTGTGGAGAGGACAGAAGCGGACAG 660

Qy    712 CGAGCCACCAAGAGAGATCAGCACACTTCATAGGGACCTTCCTTGTGTTCGGCCCTAT 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGAGCCACCAAGAAGATCAGCACACTTCATAGGGACCTTCCTTGTGTTCGGCCCTAT 720

Qy    772 GTGATCACCAGGCTAGTGGGAGCTCTTCTCCACGGGTGCCATCGGGGTCCCACTGGGGGTG 831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GTGATCACCAGGCTAGTGGGAGCTCTTCTCCACGGGTGCCATCGGGGTCCCACTGGGGGTG 780

Qy    832 CTGTCCAAGTGCTTGGCGTACAGCAAGGCGCCGCCATCCGACCCCTTTGTGTACTCCTTACTG 891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTGTCCAAGTGCTTGGCGTACAGCAAGGCGCCGCCATCCGACCCCTTTGTGTACTCCTTACTG 840

Qy    892 CGAGCACGTACCGGCAAAAGCTGCAAGGAGAGTTCTTGAACAAGGATCGGCTCCTGCCAGAACGGCTCC 951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CGACACCAGTACCGCAAAAGCTGCAAGGAGAGTTCTGAACACAGGATCGGCTCCTGCCAGAACGGCTCC 900

Qy    952 ATCCACTCCTCTGGGCCTTCACAGGCCGGACTCTCACAGCCAGACAATTCTGCCGGTGCTGAG 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ATCCACTCCTCTGGGCCTTCACAGGCCGGACTCTCACAGCCAGACAATTCTGCCGGTGCTGAG 960

Qy   1012 TGA 1014
          |||
Db    961 TGA 963
```

RESULT 2

```
LOCUS       AK036100                2802 bp    mRNA    linear   HTC 03-APR-2004
DEFINITION  Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length
            enriched library, clone:9630036A11 product:ORPHAN G PROTEIN-COUPLED
            RECEPTOR GPR26 homolog [Rattus norvegicus], full insert sequence.
ACCESSION   AK036100
```

```
FEATURES             Location/Qualifiers
     source          1..2802
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9630036A11"
                     /clone="9630036A11"
                     /tissue_type="cerebellum"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="16 days neonate"
     CDS             472..1485
                     /note="unnamed protein product; ORPHAN G PROTEIN-COUPLED
                     RECEPTOR GPR26 homolog [Rattus norvegicus] (SPTR:Q9QXI3,
                     evidence: FASTY, 99.4%ID, 100%length, match=1011)
                     putative"
                     /codon_start=1
                     /protein_id="BAC29305.1"
                     /db_xref="GI:26331150"
                     /translation="MNSWDAGLAGLLVGTIGVSLLSNGLVLVLCLLHSADIRRQPALF
                     TLNLTCGNLLCTVVNMPLTLAGVVAQRQPAGDRLLCRLAAFLDTFLAANSMLSMAALSI
                     DRWVAVVFPLSYRAKMRLRDAAFMWAYTWIHALTFPATALALSWLGFHQLYASCTLCS
                     RRPDERLRFAVFTSAFHALSFLLSFIVLCPTYLKVLARFHCKRIDVITMQTLVLLV
                     DHPSVRRCLEQKRRQRATLKKILTSFIGTFIVCFAPYVVITHLVELFSTAPIGSHWG
                     VLSKCLAYSKAASDPFVISLLRHQYRSCKELLNRIFNRSLHSVGLTGDSHSQNILP
                     VSE"
ORIGIN
```

```
Query Match         86.0%;  Score 871.6;  DB 3;  Length 2802;
Best Local Similarity    91.2%;  Pred. No. 1e-172;
Matches  925;  Conservative    0;  Mismatches  89;  Indels    0;  Gaps    0;

Qy       1 ATGGACTCGTGGGACGGCGGGCTGGCGGGCCTGGCGGGGCTTCTGGTGGGGACGATCGGC   60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     472 ATGGACTCGTGGGACGGCGGGCTGGCGGGCCTGGCGGGGCTTCTGGTGGGGACGATCGGC  531

Qy      61 CTGTCCAACGGCCTGGTGCTGGTGCTGTGCCTGCTGCACAGCGCAGATATCCGCCGCCAG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     532 CTGTCCAACGGGCTGGTGCTGGTGCTGTGCCTGCTGCACAGCGCAGATATCCGCCGCCAG  591

Qy     121 CCGGCGCTCTTCACCCTGAACCTCACGTGCGGGAACCTGCTGTGCACCGTCGTCAACATG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     592 CCGGCGCTCTTCACTCTGAACCTCACGTGTGGCAACCTGCTGTGTACCGTCGTCAACATG  651

Qy     181 CCGCTCACCCTGGCCGGCGTGGTGGCGCAGCGCCAGCCCGCCGGGGACCGCCTGTTGCGC  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     652 CCACTAACACTGGCCGGCGTGGTGGCACAGCGGCAGCCCGCCGGGGACCGCCTGTTGCGC  711

Qy     241 CTGGCTGCCTTCCTCGACACCTTCACGTGCCGGAACCTGCTGTGCAATGGCATGGCCAAG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     712 CTGGCGCCGTTCCTCGACACCTTTCTGGCCGCGAACTCCATGCTCAGCATGGCCGCGCTC  771

Qy     301 AGCATCGACCGCTGGGTGGCGTGGTCTTCCCGCTCAGCTACCGGCGCCAAGAATGCCGTC  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     772 AGCATCGACCGCTGGGTGGCGTGGTCTTCCCGCTACCGGCGCCAAGAATGCCGTC  831

Qy     361 CGCCACGCGCGCGTCATGGTGGCCTACACGTGGTGGCATCACGCTCTCCCAGCCGCC  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     832 CGAGATGCCGCCTTCATGGTGGCCTACACGTGGTGGCACCACGCTCTCCCAGCC  891

Qy     421 GCGCTCGCCCTGGGTGGCCTTCCACCAGCTGTACGCCTCGTGCACGCTGTGCAGC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     892 GCGCTCGCCCTGTCCTGGCTGGGCTTCCACCAGCTTATATGCGCTTCTGCAGC  951

Qy     481 CGGCGGCCCAGAGAGCGGCCTGCGCTTCGCCGTCTTCACCTCGGCCTTCCACGCTCTCAGC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     952 CGGCGGCCCGGACAGGAGCGGCTTTGCCGTCTTCACCAGCGCCTTCCATGCGCTCAGC  1011

Qy     541 TTCCTGCTCTTCCTTCGTTCGTTGCTTCGCTCACAGGTGCTCAAGGTGCTCAAGGTGCAGC  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1012 TTCCTGCTCTCCTTCATCGTGCTCTGCCCTACCTACCTCAAGGTGCTCAAGGTGCCGC  1071

Qy     601 TTCCATTGCAAGCGCCATCACCATGCAGACGCTGATCACCAGACGCGTTGGTGCTGTGGTGCTGGTGGACCTG  660
```

```
Db    1072 TTCCACTGCAAGCGCCATCGACGTGATCACCATGCAGACGCTGATCACCAGACGCGTGGTGTTGTTGGTGGACATA  1131
           ||||| |||||||||||||||||||||||||||||||||||||||||||| || |||| || ||||||||||| ||

Qy     661 CACCCCAGTGTGCGGAGACGCTGTCTGAGGAGCAGAAGAGAGGCGGAGAGGCAGCAGCCAGCCACC  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1132 CACCCCAGTGTGAGGAGAACGGTGTCTGGAGAGAACAGAAGCGCCAGAAGAGGCGGAGAGGCAGCAGCCAGCCACC  1191

Qy     721 AAGAAGATCAGCACCACCTTCATACGGAGACCTTCCTTGTGTTGGTGCTTGGGCCCTATGTGATACC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1192 AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGCACCCTATGTGATTACC  1251

Qy     781 AGGCTAGTGGAGCTCTTCTCCACCGCTCCCATCGGCTCCCACTGAGGGGGGGGCCTTCCAAG  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1252 AGGCTGGTGGAGCTCTTCTCCACAGCACCCATCGGCTCCCACTGGGGGGTCCTCAAG  1311

Qy     901 TACCGCAAAAGCTGCAAGGAGAGATTCTGAACACAGGCGTCCTGCACAGACGCCTCCAAGGT  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1372 TACCGGCAGGAGGCTGCAAGGAGCTCCTCGAACAGAGATCTTCAACAGACGACGGCTCCCTTCACTCT  1431

Qy     961 TCTGGCCTCACAGGCGCCGGCTCACAGCCCAGCCAGGAACAGTTCTGCGCGGGTGTCTGAGTGA  1014
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1432 GTGGGGGCCTCACAGGTGCTCACAGGTGACTCTCACACCCAGGAACATTCTGCCAGTGTCGGAATGA  1485
```

```
RESULT 3
LOCUS          AY420107                    963 bp    DNA    linear    GSS 12-DEC-2003
DEFINITION     Mus musculus HCM7113 gene, VIRTUAL TRANSCRIPT, partial sequence,
               genomic survey sequence.
ACCESSION      AY420107
VERSION        AY420107.1  GI:39776064
KEYWORDS       GSS.
SOURCE         Mus musculus (house mouse)
  ORGANISM     Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1  (bases 1 to 963)
  AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
               Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
               Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
               Adams,M.D. and Cargill,M.
  TITLE        Inferring nonneutral evolution from human-chimp-mouse orthologous
               gene trios
  JOURNAL      Science 302 (5652), 1960-1963 (2003)
  PUBMED       14671302
REFERENCE      2  (bases 1 to 963)
  AUTHORS      Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
               Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
               Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
               Adams,M.D. and Cargill,M.
  TITLE        Direct Submission
  JOURNAL      Submitted (16-Nov-2003) Celera Genomics, 45 West Gude Drive,
               Rockville, MD 20850, USA
COMMENT        This sequence was made by sequencing genomic exons and ordering them
               based on alignment.
FEATURES             Location/Qualifiers
     source          1..963
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     gene            <1..>963
                     /locus_tag="HCM7113"
ORIGIN
```

```
Query Match         81.4%;  Score 825.4;  DB 9;  Length 963;
Best Local Similarity    91.1%;  Pred. No. 4.7e-163;
Matches  877;  Conservative    0;  Mismatches  86;  Indels    0;  Gaps    0;

Qy      52 GTCTCGCCTGCTGTCAAACGGCCTGGTGCTGGTGCTGTGCCTGCTGCACAGCGCAGCGGGACATCCGC  111
```

RESULT 4
LOCUS   AY420106   666 bp   DNA   linear   GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM7113 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION   AY420106
VERSION   AY420106.1 GI:39776063
KEYWORDS   GSS.
SOURCE   Pan troglodytes (chimpanzee)
  ORGANISM Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1 (bases 1 to 666)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
  JOURNAL   Science 302 (5652), 1960-1963 (2003)
   PUBMED   14671302
REFERENCE   2 (bases 1 to 666)
  AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   This sequence was made by sequencing genomic exons and ordering them
            based on alignment.
FEATURES   Location/Qualifiers
     source   1..666
                     /organism="Pan troglodytes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9598"
     gene   <1..>666
                     /locus_tag="HCM7113"
ORIGIN

Query Match   55.0%; Score 557.2; DB 9; Length 666;
Best Local Similarity   83.9%; Pred. No. 1e-106;
Matches 559; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

Qy   1 GTGTGTCGGCTGCTTGTCGTCCAAACGGGCGCTGGTGCGTCGTGCTCGTCTGCCTCTGCAGCGCGCTGACATCCGC 60
             ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   99 CGCGGACACATCCGCCGCGGCGCCGGAGGCCCGGGGCCGCGGCGCTCTTCACCCTGAACCTCACGTGCGGGAACCT 158

Db   1 CGCGCCAGGCGGCGGCGGCGGCGCTCTTCAACCTCAGGTGGGCGGGGGAACCTGCTGCTCTCTGCAGCGCCATCCGC 60
             |||||||||||||||||||||||
Qy   1 CGCGGACACATCCGCCGCGGCGCCGGAGGCC...GCGCTCTTCACCCTGAACCTCACGTGCGGGAACCT 60

Qy   112 GTCAACACATGCCGCTCACGCTGGCGCCCGGCGTCGTGGGCGGCAGCCGGCGGCCGACCCGC 171
             |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   159 GCTGTGTCACCGTGGTCAAACATGCCCGCCGCCGCTCAGCGCTGCCGCGGCCTGCCGGCGCC 218

Db   121 GTCAACATGCCACTAACACTGGGCGCGGGCGTCGTGGCACAACGGGCACCGGCGGACCGGC 180
             |||||||||||||||||||||||||||
Qy   159 GCTGTGCACCGTGGTCAAACATGCCCGCCGCCGCTCAGCGCTGCCGCGGCCTGCCGGCGCC 120

Qy   232 CTGTCCGCGCGCTGCGGTGGCTGGCCTTCCTCGACACCACCTTCCTGGCGGCCAACTCCATGCCCATGCCCATG 291
             ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   219 GGCGCGGGCGACCGCGCCTGTGCCGCCTGCCTGCCTTCCTCTGGACACCTTCCTGCTGCCAACTC 278

Db   181 CTGTCCGCGCTGGCCCGGCCGCCTTCCTCGACACCTTTCTGGCGGCCCAACTCCAATGCTCCAGCATG 240
             |||||||||||||||||||||
Qy   219 GGCGCGGGCGACCGCGCCTGTGCCGCCTGCCTGCCTTCCTCTGGACACCTTCCTGCTGCCAACTC 180

Qy   292 GCCGCGGCGCTCAGCGCATGCAATCGACCGCTGGGGTGGGCCGGCCGTGGGGTCTTCCCGCTGAGCTAC 351
             |||||||||||||||||||||||||||||||
Db   279 CATGCTCAGCATGGCCGCGCGTCAGCATCGACCGCTGGGGTGGGTCTCTTCCCGCTGAG 338

Db   241 GCCCGGCGCTCAGCACTCGACCGCTGGGGTGGGCTCGTTCGGCTGTGGGTTTCTTCCCGCTGAGCTACGCTGCGCAAG 300
             |||||||||||||||||||
Qy   279 CATGTCAGCATGGCCGCGCGTCAGCATCGACCGCTGGGGTGGGTCTCTTCCCGCTGAG 240

Qy   352 ATGGCGCCCTCCGCGACGCGGCGGGCGGCCTCATGGTGGCCTGCTACACGTGGCCTCACGTGGCCTCACCTTC 411
             ||||||||||||||||||||||||
Db   339 CTACCGGGGCCAAGATGCGCCTCCGCGCCGACGCGGCGCCGCGCCGCCATGGTGGCTGCCTACACGTGGCCTGCA 398

Db   301 ATGGCGCCCTCCGAGATGCGCCGCTTCATGGTGGCCTACACGTGGCCTCACGTGGCCTCACCTTC 360
             ||||||||||||||||||||||
Qy   339 CTACCGGGGCCAAGATGCGCCTCCGCGCCGACGCGGCGCCGCGCCGCCATGGTGGCTGCCTACACGTGGCCTGCA 300

Qy   412 CCAGCGGCGCGGCCTGTGCTCCTGCTGGGGGTGCTGGCTTCCTCTCCACCCTGCTCCGTCTGCTCGTACCCCTCCGTGCGCACG 471
             ||||||||||||||||||
Db   399 CGCGCTCACCTTCCCAGCGCCCGGCCTGTGCTGCGCCTGCTGGGGGCGCGCTTCGGCCTTCGGCCTTCGGCCAGCAGTA 458

Db   361 CCGGCGCACCGCGCTCGGCCGTGTCCTGCCACTCCACCAGCTATATGCCTCGTGTGCACA 420
             |||||||||||||||||||
Qy   399 CGCGCTCACCTTCCCAGCGCCCGGCCTGTGCTGCGCCTGCTGGGGGCGCGCTTCGGCCTTCGGCCTTCGGCCAGCAGTA 300

Qy   472 CTGTGCCAGCGCCGCTTCCATTGCAAGCGCATCGAGCGTGATCACATCGACGAGCTGCCACGCACCTTCCAC 531
             |||||||||||||||||||||||||
Db   459 CGCCTCGTGCACGCGTGTGCACGCGGCCCAGACAGCGAGCGGCGCTGCGCTTCGGCGTTCGGCGTCTTCAC 518

Db   421 CTGTGCCAGCGCCGGGGACGCATGGCTCGGGGTCCTGGCGTGCTCTCTCCACCGACCGGCCTTCCAT 480
             |||||||||||||||||
Qy   459 CGCCTCGTGCACGCGTGTGCACGCGGCCCAGACAGCGAGCGGCGCTGCGCTTCGGCGTTCGGCGTCTTCAC 420

Qy   532 GCTCTCTCAGCTTCCTCCTTCTTCGTCGTGCTGCTTCACGTGACCTACCTCAAGGTGCTCAAG 591
             ||||||||||||||||||||
Db   519 TGGCGCCTTCCACGCTCTCAGCTTCCTCTGCTGCTTCCTCGTCGTGCTGCTTGCCGCTTGCTCGTGCTGCCACGTGTGACCT 578

Db   481 GCGCTCTCAGCTTCCTCTCTTTCATCGTGCTGCTTCACGTGACCTACCTCAAGGTGCTGCTCAAG 540
             ||||||||||||||
Qy   519 TGGCGCCTTCCACGCTCTCAGCTTCCTCTGCTGCTTCCTCGTCGTGCTGCTTGCCGCTTGCTCGTGCTGCCACGTGTGACCT 480

Qy   592 GTGGGCCCGCTTCCATTGCAAGCGCATCGAGCGTGATCACATCGACGAGCTGCCAACTCN 651
             |||||||||||||||||
Db   579 CAAGGTGCTCAAGGTGGCCCGCCCGCTTCCATTGCAAGCGCCATCGACACGTGATCACCACCATCGACGAC 638

Db   541 GTGGGCCCGCTTCCACTGCAAGCGCATCGAGTCGATCACACGCTTGTGCTGGTTG 600

Qy   652 GTGGACCTGCACCCAGTGTGCGGAAGCGTGTCTGGAGGAGCAGAAGCCGGAGCAGAG 711

Db   601 GTGGGACATACACCCCAGTGTTGAGGGAAACGGTGTCTGGAGGAACAGAAGCCGGAGCAGAG 660

Qy   712 CGAGCCCACCAAGAGAGATCAGCACTCAGCACTTCATAGGGAACTTGTGCTTCGCGCCCTAT 771

Db   661 CGTGCCGCACCAAGAGAAGATCAGCACCTTCATGAGGGTGGCTTTGCACCCTAT 720

Qy   772 GTGATCACCAGGCTAGTAGTGGAGCTGCATGCATGCATGCAGTGATCACATTGCAACTGTGGCTG 831

Db   721 GTGATTACCAGGCTGGGTGGAACTCTTCTCCACAGACAGCCCAITGGCTCACTGCAGTG 780

Qy   832 CTGTGTCAAGTGCTTGGCGTACAGCAAGAGGCCGCATCCGACCCCCTTTGTGTACTCTTACTG 891

Db   781 CTGTGCCAAGTGCTTGGCGTACAGCAGCAAGAGCCCGACCCCTTCGGTGTATTCCTTGCTG 840

Qy   892 CGACACCCAGGTACCGCGCAAAAGGTGCAAGAGAGTGTCAAGGAGGACTGCAAAGAGAACCTCC 951

Db   841 CGACACCAAATACCGCAGGAGACTCTTCGAACAGGAGTCTTCAACAGAGACGCTCC 900

Qy   952 ATCCACTCCTTCTGGGCCTCACAGGCGGACTCTCACAGCCAGACATTCTGCCGCGGTGTCTGAG 1011

Db   901 CTTCACTCCTGTGGGGCCTCACAGCCAGCAGGTGGACTCTCACAGGAACATTCTGCCCAGTGGTCGGGAA 960

Qy   1012 TGA 1014
             |||
Db   961 TGA 963

RESULT 5

LOCUS   CO934052   741 bp   mRNA   linear   EST 16-AUG-2004
DEFINITION   AGENCOURT_30488535 NIH_MGC_145 Homo sapiens cDNA clone
    IMAGE:7262026 5', mRNA sequence.
ACCESSION   CO934052
VERSION   CO934052.1   GI:51288729
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 741)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
    Office of Cancer Genomics
    National Cancer Institute / NIH
    Bldg. 31 Rm10A07 Bethesda, MD 20892
    Email: cgapbs-r@mail.nih.gov
    Tissue Procurement: GPCR Consortium
    cDNA Library Preparation: GPCR Consortium
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: Agencourt Bioscience Corporation
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at:
    http://image.llnl.gov
    Plate: IRB16 row: c column: 08
    High quality sequence stop: 607.
FEATURES   Location/Qualifiers
  source   1..741
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:7262026"
    /tissue_type="mixed"
    /clone_lib="NIH_MGC_145"
    /lab_host="DH10B"
    /note="Vector: pcDNA3.1; Site_1: varies by clone; Site_2:
    varies by clone; ORFs were PCR-amplified and cloned into
    pcDNA3.1 by the GPCR Consortium. Cloning sites vary by
    clone and include the following: 5'-EcoRV-XmnI/XhoI-3',
    5'-EcoRV-XmnI/NotI-3', EcoRV (TA cloned, non-directional).
    For information about which gene each clones represents,.
    please visit our anonymous ftp site at
    ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
    a Note: this is a NIH_MGC Library."

ORIGIN

Query Match   53.1%;   Score 538.8;   DB 7;   Length 741;
Best Local Similarity   95.9%;
Matches 585; Conservative   0; Mismatches 22; Indels 3; Gaps 3;

Qy   1 ATGAACTCGTGGGAACGGACGGGCGCCTGGCGGGGGGGGCTACTGGTGGGGCACCATAGGAGGGCGTCGGCTG   60
       |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16 ATGAACTCGTGGGAACGGACGGGCGCCTGGCGGGGGGGGCTACTGGTGGGGCACCATAGGAGGGCGTCGGCTG   75

Qy   481 CAAGGTGGTGCTCAAGGTGCCATCCGCTTCCATTGCAATGGCCATCGGACGGTGATCACCATGTGCAGAC   540
       ||||||
Db   541 GCTCGTGCTGCTGGTGGACCTGCACCGGGGAACCGCTGGTGTCTGGAGGAGGAGCAGAA   600

Qy   541 GCTCGTGCTGCTGGTGGACCTGCACCGGGGAACCGCTGGTGTCTGGAGGAGGAGCAGAA   600
       ||||||||||
Db   639 GCTGGTGCTGCTGGTGGACCTGCACCGGGGAACGGCTGGTGTCTGGAGGAGGAGCAGAA   698

Qy   601 GCGGACGGCGACGAGCAGCAGCAGCCAAGAAGGATCNNNCACGTTCATAGGAGGACCTTCCTTGTGTG   660
       ||||||||
Db   699 GCGGACGGCGACGAGCAGCAGCCACCAAGAAGGATCAGCACACCTTCATAGGAGGACCTTCCTTGTGTG   758

Qy   661 CTTCGC   666
       ||||||
Db   759 CTTCGC   764

RESULT 6

LOCUS   CO934490   736 bp   mRNA   linear   EST 16-AUG-2004
DEFINITION   AGENCOURT_30488533 NIH_MGC_145 Homo sapiens cDNA clone
    IMAGE:7211818 5', mRNA sequence.
ACCESSION   CO934490
VERSION   CO934490.1   GI:51289167
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 736)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
    Office of Cancer Genomics
    National Cancer Institute / NIH
    Bldg. 31 Rm10A07 Bethesda, MD 20892
    Email: cgapbs-r@mail.nih.gov
    Tissue Procurement: GPCR Consortium
    cDNA Library Preparation: GPCR Consortium
    cDNA Library Arrayed by: Agencourt Bioscience Corporation
    DNA Sequencing by: Agencourt Bioscience Corporation
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at:
    http://image.llnl.gov
    Plate: IRB15 row: c column: 08
    High quality sequence stop: 600.
FEATURES   Location/Qualifiers
  source   1..736

Qy   61 CTGTCCAACGCGTGGTTGGTTGCTTCGTTCGCTTGCCTTGCACAGGGCGGGACATCCGGCGCCAGGCG   120
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76 CTGTCCAACGCGTGGTTGGTTGCTTCGTTCGCTTGCCTTGCACAGGGCGGGACATCCGGCGCCAGGCG   135

Qy   121 CCGGACGCGTCTCACCCTGAACCTCAGCTGCGGGAACCTGCTGTGCACCGTGGTCAACATG   180
       ||||||||||||||||||||||||||||||||
Db   136 CCGGGCGTCTTCACCCTGAACCTCAGCTGCGGGAACCTGCTGTGCACCGTGGTCAACATG   195

Qy   181 CCGCTCACGCTGGCCGGCGGCGGGGGGGCGGCAGGGAGCGGGAGGCAGCAGCGCCGCTGTCCGGC   240
       ||||||||||||||||||
Db   196 CCGCTCACGCTGGCCGGCGGCGGGGGGGCGGCAGGGAGCGGGAGGCAGCAGCGCCGCTGTCCGGC   255

Qy   241 CTGGCTGCCTTCCTCGACACCTTCCTGGCTGGCCAACTCCATGCTCAGCATGGCCGGCCTC   300
       ||||||||||||||||
Db   256 CTGGCTGCCTTCCTCGACACCTTCCTGGCTGGCCAACTCCATGCTCAGCATGGCCGGCCTC   315

Qy   301 AGCATCGACCCGCTGGGTGGGCGTGGTCTTCCGGCTGCCACCTTCCCAGCAGCCGCC   360
       ||||||||||||||
Db   316 AGCATCGACCCGCTGGGTGGGCGTGGTCTTCCGGCTGCCACCTTCCCAGCAGCCGCC   375

Qy   361 CCGGACGCGGCGGGGCGGCTCATGGGTGGCCTACACGTGGGCTGCCACGCGGCTCACCTTCCCAGCAGCC   420
       ||||||||
Db   376 CCGGACCGGGCGGCTCATGGGTGGCCTACACGTGGGCTGCCACGCGGCTCACCTTCCCAGCAGCC   435

Qy   421 GCGGCTCGGCCCTGTCCTGGGGCGTCCGGCTTCCACCAGCTGTACGGCCTGGTCGCAGC   480
       ||||||||||||
Db   436 GCGGCTCGGCCCTGTCGGGCTTCGGCTTCCACCAGCTGTACGGCCTGGTCGCAGC   495

Qy   481 CGGCGGCCAGGACGAGGCGGCCCTGGCGCTTCGCGCCGTCTTCCACGCGCTCGTGCAGC   540
       ||||||||||||
Db   496 CGGCGGCCGGACGAGGCGGCCCTGGCGCTTCGCGCCGTCTTCCACGTGGCGGC-TCCAGCGCTCTCAGC   554

Qy   541 TTCCTGCGTCTCCTTCGTCGTCGTCCTGCTGCACGTACCTACCTCAAGGTGTCTCAAGGTGGCCGC   600
       ||||||||
Db   555 -TCCTCGTCTCCTT-GTCGTCGTCCTGCACGTACTCACGGTGCTCAAGTGCCGGCGCTTCA   612

Qy   601 TTCCAATTGCA   610
       ||| |||
Db   613 TTGCAGCGCA   622

ORIGIN
FEATURES   Location/Qualifiers
  source   1..736

ORGANISM   Homo sapiens (human)
SOURCE   Homo sapiens
KEYWORDS   EST.
VERSION   CO934461.1 GI:5289138
ACCESSION   CO934461
DEFINITION IMAGE:7211818 3', mRNA sequence.
LOCUS   CO934461   730 bp   mRNA   linear   EST 16-AUG-2004
            AGENCOURT_30698355 NIH MGC 145 Homo sapiens cDNA clone

RESULT 7

CO934461/c

Qy   Matches 560; Conservative   0; Mismatches   8; Indels   4; Gaps   3;
Best Local Similarity   97.9%; Pred. No. 2.9e-99;
Query Match

ORIGIN

Qy   1 ATGAACTCGTGGGACGCGGCGGGCGGCTACTGGTGGGCACGATAGGGGCGTCTCGCTG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   16 ATGAACTCGTGGGACGCGGCGGGCGGCTACTGGTGGGCACGATAGGGGCGTCTCGCTG   75

Qy   61 CTGTCCAACGCGCTGGGTGCTGCTCCTCTGCCCTGCGACATCCGCGCGCGCCAGGCG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   76 CTGTCCAACGCGCTGGGTGCTGCTCCTCTGCCCTGCGACATCCGCGCGCGCCAGGCG 135

Qy   121 CCGGGGCGCTCTTCACCCTGAACCTCACGTGGGGGGAACCTGGTGTGGTCAACATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   136 CCGGGGCCTCTTCACCCTGAACCTCACGTGGGGGGAACCTGGTGTGGTCAACATG 195

Qy   181 CCGCTCACGTGGCCGGCGTCGTGGCGGCCCAGCCGGCGGAACCCCCGTTGTGCCGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   196 CCGCTCACGGTGGCCGGCGTGGCGCCAGCCGGCGGAACCCCCGTTGTGCCGC 255

Qy   241 CTGGCTGCCTTCCTCCGACACCTTCCTGGCGCTGCCAACTCCATGCTCAGCATGGCCCGCGGCGCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   256 CTGGCTGCCTTCCTCCGACACCTTCCTGGCGCTGCCAACTCCATGCTCAGCATGGCCCGCGGCGCTC 315

Qy   301 AGCATCGACCGCTGGGGTGGTGGGGCCGGTGGTCTTCCCGGTGAGCTAGCTACCGGCCAAGATGCGGCCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   316 AGCATCGACCGCTGGGGTGGTGGGGCCGGTGGTCTTCCCGGTGAGCTAGCTACCGGCCAAGATGCGGCCTC 375

Qy   361 CGCGACGCGCGCGCCGCCTCATGGGTGGCCTGCGTGCCACCGCGCTCACCTTCCCAGCCGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   376 CGCGACGCGCGCGCCGCTCATGGGTGGCCTGCGTGCCACCGCGCTCACCTTCCCAGCCGCC 435

Qy   421 GCGCTCGCCCCTGTCCTGTGGCCTGTGGCCCTACACGTGTGCCACGGCGCTGTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   436 GCGCTCGCCCCTGTTCCTGGCCTCGTTGCACCCTACACGTGTGCCACGGCGCTGTGCAGC 495

Qy   481 CGGGCGCCCAGACGAGCGGCCTCGCGCTCGCGGTCTTCGCGGTCCTTGCCGGTCCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   496 CGGGCGGCCGGACGAGCGGCTTCGCGGTTCNCGGTCTT-ACTGGCGCCTT-CAGCTCTCAGC 553

Qy   541 TTCCTGCGCTCTCCTTGGTCGTGTCTGCTGCGCAC 572
          ||||||||||||||||||||||||||||||||
Db   554 TTC--TGCTCTCTTGGCGCTGCTCTGCTGCGCAC 583

```
Qy    811  ATCGGCTCCCACTGGGGGGGTCGTCGTTGGCGTACAGCAGGCCGAAGCCCGCATCCGAC 870
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    272  ATCGGCTCCCACTGGGGGGGTGCTGTCCAAGTGCTTGGCGTACAGCAGGCCGCATCCGAC 213

Qy    871  CCCTTTGTGTACTCCTTACTCGGACACCAGTACCGCAAAAGCTGCAAGGAGGAGATTCTGAAC 930
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    212  CCCTTTGTGTACTCCTTACTCGGACACCAGTACCGCAAAAGCTGCAAGGAGGATTCTGAAC 153

Qy    931  AGGCTCCTGCACAGACGCTCCATCCACACTCCTTCTGGCCTCACAGGGGACTCTCAGCCAG 990
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    152  AGGCTCCTGCACAGACGGCTCCATCCACTCCTTCTGGCCTCACAGGGCCACTCTCACAGCCAG 93

Qy    991  AACATTCTGCCGGTGTCTGAGTGA 1014
           ||||||||||||||||||||||||
Db    92   AACATTCTGCCGGTGTCTGAGTGA 69
```

RESULT 8

LOCUS   AK042755   1211 bp   mRNA   linear   HTC 03-APR-2004

DEFINITION Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
            enriched library, clone:A730021D22 product:ORPHAN G PROTEIN-COUPLED
            RECEPTOR GPR26 homolog [Rattus norvegicus], full insert sequence.

ACCESSION   AK042755

VERSION   AK042755.1 GI:26335306

KEYWORDS   HTC; CAP trapper.

SOURCE   Mus musculus (house mouse)

  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
   PUBMED   10349636

REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
   PUBMED   11042159

REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
   MEDLINE 20530913
   PUBMED   11076861

REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)

REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)

REFERENCE   6 (bases 1 to 1211)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saito,H., Sakai,J., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)

COMMENT   cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
            URL:http://fantom.gsc.riken.jp/.

FEATURES   Location/Qualifiers
     source   1..1211
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:A730021D22"
                     /db_xref="taxon:10090"
                     /clone="A730021D22"
                     /tissue_type="cerebellum"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="7 days neonate"
     CDS   <1..584
                     /note="unnamed protein product; ORPHAN G PROTEIN-COUPLED
                     RECEPTOR GPR26 homolog [Rattus norvegicus] (SPTR|Q9QXI3,
                     evidence: FASTY, 99.4%ID, 100%length, match=1011)
                     putative"
                     /codon_start=3
                     /protein_id="BAC31354.1"
                     /db_xref="GI:26335307"
                     /translation="SWLGFHQLYASCTLCSRRPDERLRFAVFTSAFHALSFLLSFPIVL
                     CFTLKVLKVARFHCKRIDVITWQTLVLLVDIHPSVRERCLEEQRRRQRATKKISTF
                     IGFTLVCFAPYVITRLVELFSTAPIGSHWGVLSKCLAVSKAASDPFVYSLLRHQYRRS
                     CKELLNRIFNRRSLHSVGLTGDSHSQNILPVSE"

ORIGIN
```

Query Match         48.5%;    Score 491.8;   DB 3;   Length 1211;
Best Local Similarity    90.2%;    Pred. No. 6e-93;
Matches 526;  Conservative  0;   Mismatches 57;   Indels 0;   Gaps 0;

```
Qy    432  GTCGTTGGCGTACAGCAGCCTTCCACCAGCTGTACGCCTCGTCGTCACGCGTGTGAGCCGGGCGGGGCCAGA 491
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2    GTCCTGGCTCGGCTTCCACCAGCTATATGCCTCGTTCGTCGTCACACTGTCGCAGCCGCCGGGA 61

Qy    492  CGAGCGCCCTCGGCTTCGCCCTCTTCACTCGCGCCTTCCAGCTTCCTGCTCCTCTCTC 551
           ||||||||||||||||| || |||||||||||||||||||||||||||||||||||||
Db    62   CGAGCGCCCTCGGCGTTTGCTGTCTTCACCAGGCCTTCCATCGCCTCAGCTTCCTGCTCTC 121

Qy    552  CTTCGTCGTCGTTCGTCGTGTGCACGTACCTCAAGGTGCTCAAGGTGGCCCGCGCCTTCCAATTGCAA 611
           || |||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db    122  CTTCAATCGTGTCGTCGTTCACGTACCTCACGGTGCTCAAGGTGGCCCGCGCTTCCACTGCAA 181

Qy    612  GCGCATCGACGTGATCACCATGCAGACGCTGGTGCTGCTGGTGGACCTGCACCCCAGTGT 671
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182  GCGCATCGACGTGATCACCATGCAGACGCTTGTTGTCTTGGTGGACATACACCCCAGTGT 241

Qy    672  GCGGGAACGCTGTTCTGGAGGAGCAGAAGCGGCAGCGGCAGCCACAAAGAAGAGATCAG 731
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    242  GAGGGAACGCTGTCTGGAGGAACAGAAGCGGAGGCGGCAGCAGCGTCGCCACCAAGAAGATCAG 301
```

RESULT 9

```
Qy   732  CACCTTCATAGGGGACCTTCCTGCGGCCCCCTATGTGATCACCCAGGCTAGTGGA  791
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   302  CACCTTCATAGGGGACCTTCCTGCGGCCCCCTATGTGATCACCCAGGCTAGTGGA  361

Qy   792  GCTCTTCTCCACGGTGCCCATCGGCCCACTGGGGGGTGCTGTGTCCAAGTGCTTGGGGTA  851
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   362  ACTCTTCTCCACAGCACCCATTGGCTCTCCACTGGGGGAGTGCTGTCCAAGTGCTTGGGCTA  421

Qy   852  CAGCAAGGCCGCCATCCAGACCACCCCTTTGTTGTACTCCTTACTGCGACCAGTACCCAAAAG  911
          ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   422  CAGCAAGGCCGCCGTTCTGACCCCCTTCGTGTATTCCTTGCTGCGACCAGTACCCGAGGAG  481

Qy   912  CTGCAAGGGAGATTCTGAACAGGCTCCTGACACACCAGTACTCCTTCTGGCCTCAC  971
          |||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db   482  CTGCAAGGGAGCTCCTTGAACAGGATCTTCAACAGACGCGTCCCCTTCACTCTGTGGGCCTCAC  541

Qy   972  AGGCCGACTCTCTCACACCCAGACCACATTCTGCCGGGTGTCTGAGTGA  1014
          ||| ||||||||||||||||||||||||||||||||||||||||||
Db   542  AGGTGACTCTCACACCCAGACGAACATTCTGCCAGTGTCGGGAATGA  584
```

ORIGIN   program coordinator."

```
Query Match
Best Local Similarity   47.7%;
Score 483.4;  DB 6;  Length 861;
Best Local Similarity   90.5%;  Pred. No. 3.4e-91;
Matches   514;  Conservative    0;  Mismatches   54;  Indels    0;  Gaps    0;

Qy     1  ATGAACTCGTGGGACGGGCGCCTGGCGCGGGGCGCTACTGGTGGCGCCACGGATGGGCGGTCTCGCTG  60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   293  ATGAACTCGTGGGACGGGCGCGGGGCGCCTGGCGCGGGGCGCTACTGGTGGCGCCACGGATGGGCGGTCTCGCTG  352

Qy    61  CTGTCCAACCGCGGCTGGTGCTGTCGCCGCTGTGGCGCACAGCGCGGACATCCGCGCGGTGTCGCTG  120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   353  CTGTCCAACCGGCGGCTGGTGCTGTCGCCGCTCGCGCACAGCGCGGACATCCGCGCGGTGTCGCTG  412

Qy   121  CCGGGCGGCTCTTCACCCTGAACCTCACGTTGGGGGGAACCTTGGTCTGTTGGTCAACATG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   413  CCGGGCGGCTCTTCACTCTCACGTTGGGGGGAACCTTGGTTACCGTGGTCAACATG  472

Qy   181  CCGGCTCACGCTGCGGGCGGTGGCCGGGGGCCAGCGGCGCGGGGGGGGGACCCCGTGCCGC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   473  CCACTAACACTGGGCGGCGGCGTCGTGGCCACAACCGGGGCGGGGGGGGGACCCCGTGCCGC  532

Qy   241  CTGGGCTGCCTTCCTCGGCTGCCCAACCTTCCTGGCCAACTCCATGGCTCCGGGCTCG  300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   533  CTGGGCCGCCTTCCTCGACACCTTTCTGCGGCCCAACTCCATGGCATGGCCCGGCTC  592

Qy   301  AGCATCGACCGGCTGGGTGGCGGTGGCGTCTTCCCGGTGAGCTACCGGGCCAAGGATGGCCCGGCCTC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   593  ANCATCGACCGGCTGGGTGGCGGTGGTGTCTTTCCGCTGAGCTACCGGTCCAAGGATGGCCCCTC  652

Qy   361  CCGCGAGCGGCGCCAGACGAGGCGCCATTGGTGGCCTACACGTGGCGCGGGCTCACCCTTCCCAGCGCGCC  420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   653  CGAGACGGCGCCGCCGGCCTTCATGGTGGCCTACACGTGGCGCGGTCACCTTTCCGGCCAC  712

Qy   421  GCGCTGCGCCTGTCCTGGCGGGCCCTTCCACCACAGCGTACCCCTCCTGGCGCGTGTGCAGC  480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   713  GCGCTGCGCCCTGTCCTGGCGGTCGGGCTTCCACCAGCTATATGGCTCGTGCACACTGTGCAGC  772

Qy   481  CGGCGGCCAGACGAGGCGCGGCGTTCGCCGGTCTTCACGCGCTCAGC  540
          ||||||||||||||||||||||||||||||||||||||||||||||
Db   773  CGGGCGCGGCCGGACGACGCGGCGCGCTTTGCGTGTNNCTCACAGCGCCTTCCAGCGGCTCAGC  832

Qy   541  TTCCTGCTCTCCTTCGTCGTGCTCTGCT  568
          ||||||||||||||||||||||||||||
Db   833  TTCCTGCTCTCCTTCATCGTGCTCTGCT  860
```

RESULT 10

```
          The Institute of Physical and Chemical Research (RIKEN)
          1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          Tel: 81-45-503-9222
          Fax: 81-45-503-9216
          Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
  Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
  Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new
  genes. Genome Res. 10 (10), 1617-1630 (2000)
  Watahiki,A., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
  wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
  Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
  and Hayashizaki,Y.
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multicapillary sequencer. Genome Res.
  10 (11), 1757-1771 (2000)
  Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
  Sugahara,Y. and Hayashizaki,Y.
  Computer-based methods for the mouse full-length cDNA
  encyclopedia: real-time sequence clustering for construction of a
  nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
  Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A.,
  Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K.,
  Arakawa,T., Ishii,Y. and Hayashizaki,Y.
  Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
  Func. Genomics 2 pre, L72-L86 (2001
  Please visit our web site (http://genome.gsc.riken.go.jp) for
  further details.
  e mouse tissues.
FEATURES             Location/Qualifiers
  source           1..703
                   /organism="Mus musculus"
                   /mol_type="mRNA"
                   /db_xref="taxon:10090"
                   /clone="A730021D22"
                   /tissue_type="cerebellum"
                   /dev_stage="7 days neonate"
                   /lab_host="DH10B"
                   /clone_lib="RIKEN full-length enriched, 7 days neonate
                   cerebellum"
                   /note="Site 1: SalI; Site 2: BamHI; cDNA library was
                   prepared and sequenced in Mouse Genome Encyclopedia
                   Project of Genome Exploration Research Group in Riken
                   Genomic Sciences Center and Genome Science Laboratory in
                   RIKEN. Division of Experimental Animal Research in Riken
                   contributed to prepare mouse tissues. 1st strand cDNA was
                   primed with a primer [5'
                   GAGAGAGAGAGAGAGAGAGAGAGCTCTTTTTTTTTTTTTTTTTVN 3'],
                   prepared by using trehalose thermo-activated reverse
                   transcriptase and subsequently enriched for full-length by
                   cap-trapper. cDNA went through one round of normalization
                   to Rot = 20.0 and subtraction to Rot = 459.0. Second
                   strand cDNA was prepared with the primer adapter of
                   sequence [5' GAGAGAGAGAGAGAGAGAGAGAGAGAGAGCCCCCCCCCCCCCC
                   3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                   modified pBluescript KS(+) after bulk excision from Lambda
                   FLC I."
ORIGIN
  Query Match         47.4%;    Score 480.8;    DB 2;    Length 703;
  Best Local Similarity   90.1%;    Pred. No. 1.2e-90;
  Matches  526;    Conservative   0;    Mismatches  57;    Indels   1;    Gaps   1;

Qy        432 GTCCTGGCGTCGGCTTCCACCAGCTGTCACGCCTTCGTGTGCACGGCGGCGGCGGCCAGA 491
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GTCCTGGCGTCGGCTTCCACCAGCTATATGCCTTCGTGTGCACACGTGTTCGAGCGCCGGA  61

Qy        492 CGAGCGCCCTGCGCTTTGCTTCGCCTTCCAGGCTTCAGCTTCCTGCCTCTC 551
              ||||||||||||||||||||||| ||||||||| ||||| ||||||
Db         62 CGAGCGCCCTGCGCTTTGCTGTCTTCACCAGGCGCCTTCAGCTTCCTGCCTCTC 121

Qy        552 CTTCGTCGTTCGTGCTGCTGCACGTACCTCAAGGTGGCGCCCGCTTCCATTGCAA 611
```

```
RSULT 11
BH118705
BH118705/c
LOCUS         BH118705              537 bp    DNA    linear    GSS 19-JUL-2001
DEFINITION    RPCI-24-359A9.TV RPCI-24 Mus musculus genomic clone RPCI-24-359A9,
              genomic survey sequence.
ACCESSION     BH118705
VERSION       BH118705.1  GI:14961260
KEYWORDS      GSS.
SOURCE        Mus musculus (house mouse)
  ORGANISM    Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1  (bases 1 to 537)
  AUTHORS     Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
              Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
              Russell,D., de Jong,P. and Fraser,C.M.
  TITLE       Mouse BAC End Sequences from Library RPCI-24
  JOURNAL     Unpublished (1999)
COMMENT       Other_GSSs: RPCI-24-359A9.TJB
              Contact: Shaying Zhao
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: szhao@tigr.org
              Clones are derived from the mouse BAC library RPCI-24. For BAC
              library availability, please contact Pieter de Jong
              (pdejong@mail.chori.org). Clones may be purchased from BACPAC
              Resources (http://www.chori.org/bacpac/ordering/frame.htm). BAC end
              page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
              Plate: 359   row: A   column: 9
              Seq primer: T7
              Class: BAC ends.
FEATURES             Location/Qualifiers
  source           1..537
                   /organism="Mus musculus"
                   /mol_type="genomic DNA"
                   /strain="C57BL/6J"
                   /db_xref="taxon:10090"
                   /clone="RPCI-24-359A9"

Db        122 CTTCATCGTGCTGCTGCTTCACGTACCTCAGTGTGCTGGTGGACCTGCCACCCCAGTGT 181
Qy        612 GCGGCATCGACGTGATCACCATGCAGAGCGCTGGGGTGCTCGGGTGGGTGGCCCGGCCAGTGT 671
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 GCGGCATCGACGTGATCACCATGCAGAGCGGCTTGTTGGTTGGTGGCATACACCCCAGTGT 241
Qy        672 GCGGGAACGCGTGTCTCGAGGAGCAGGAAGCGGAGGCGGAGCAGGCCACCAGGAAGGATCAG 731
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 GAGGGAACGCGTGTCTCGAGGAACAGGAAGCGGAGGCGGAGCAGGCCGCCAGGAAGATCAG 301
Qy        732 CACCTTCATAGGGACCACTTCCTTGTGTGTGCTTCGCGCCCCTATGTCATACACGGCTAGTGGA 791
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CACCTTCATAGGGACCACTTCCTTGTGTGTGCTTCGCACCCTATGTCATTACACGGCTGGTGGA 361
Qy        792 GCTCTTCTTCACGGTGCCCATCGGGTGCCCACTGGGGGGTGGTGTTCCAAGTGCTTGGGTA 851
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 ACTCTTCTTCCACAGCACCCCATTGGCTTCCACTGGGGAGTGCTTGGTTCCAAGTGCTTGGGCTA 421
Qy        852 CAGCAAGGCCCGCATCCGACCCCTTCTTTGTGTACT--CCTTACTCGCGACACCAGTACCGGCAAA 910
              ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db        422 CAGCAAGGCCCGCGCTTCTTTGACCCCCTCGTGTATTCCCTTGCTCGCGACACCAAATACCGGAGGA 481
Qy        911 GCTGCAAGGAGGATTCTGAACAGGCTCCTCGCACAGACGCCTCCATCCACTCCTTCTGGCCCTCA 970
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        482 GCTGCAAGGAGCTCCTGAACAGGAGATCTTCAACAGACGCTCCCTTCACTCGTGTGGGGCCTCA 541
Qy        971 CAGGGGACTCTCCACAGCCAGAACATTCTGCCGGGTGTCTGACTGA 1014
              ||| |||||| |||| ||||||||||||||| ||||||||||
Db        542 CAGGTGACTCTCCACAGCCAGAACATTCTGCCAGTGTCGGAATGA 585
```

Left column

```
ORIGIN
          /sex="Male"
          /cell_type="Spleen/Brain"
          /clone_lib="RPCI-24"
          /note="Vector: pTARBAC1; Site_1: BamH1; Site_2: BamH1; The
          RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
          library was cloned in the pTARBAC1 cloning vector at the
          BamHI sites using MboI partially digested male C57BL/6J
          DNA."

Query Match    38.7%;  Score 392.4;  DB 8;  Length 537;
Best Local Similarity   90.9%;  Pred. No. 4.3e-72;
Matches   429;  Conservative   0;  Mismatches  41;  Indels  2;  Gaps  1;

Qy  199 GTCGTGGGCCGAGCGGCCGCAGCCGGCGGGGGGGACC--GGCTGGTGTCGCTGGGTTCCTTCCTCG 256
        ||||||||||  ||| ||  ||||| ||||||||||   ||||||||||||  ||||||||||||
Db  537 GTCGTGGGCCCAACGGCACCATCCGGCGGGGGGGTGCCGGGGCGGCCGGGCGGCCTTCCTCG 478

Qy  257 ACACCTTCCTTGGCGCTGCCACACTTCCATGCTCAGCATGGCACGGCGGCGCTCAGCCATCGACCGGCTGGG 316
        |||||||  |||| ||  ||||||| ||||| |||||||||||  |||| ||||||| |||||||||||||
Db  477 ACACCTTTCTTGGGCGCCACACTCCCATTGCTCAGCATGGCCCCGGCCGCTCAGCGATCGACCGGCTGGG 418

Qy  317 TGGCCGGGTCTTCCCGCTGAGCTACCGCGGGGGCCAAGATGCGCCTCCGCGAGCGGGCGGCTCA 376
        |||| ||| |||| ||| ||||||||| | ||||||||||| ||||||||||||| ||| ||||
Db  417 TGGCTGGGTTCTTCCGGTGAGCTACCGCTCGCCCAAGATGCCCTCCGAGAGATGCCGGCCTTCA 358

Qy  377 TGGTTGGGGCCTACACGTGGGCGTGCCAGCGGGGCTCACCTTCCCACGCGCGCTCGGCCTTGTCCTT 436
        ||||| ||||||||| |||| |||||||||||||||||||| ||  ||||||||| |||||| |||
Db  357 TGGTTGGGGCCTACACGTGGGCTCGCAGCGGGGGCTCACCTTCCCCACGCGCACTGGCCCCCTGTCCTT 298

Qy  437 GGCTCGGCTTCCACCCAGCAGTGTACGGCCTTCGTTGCCAGCGGCCAGACGAGC 496
        ||||| ||| ||||| | |||| ||||||||||| ||||||||||||||||||
Db  357 GGCTCGGCTTCCACACAGCGTATATGCCCTCGTGCACACTCGCGCCGGCGAGC 238

Qy  497 GCCTGGCTTGCGCCGGTGCTCTCACTGGGCGCCTCTCAGCTTCCTGCTCTCCTTCG 556
        |||| ||||||||| ||| ||| ||||||| || ||||| || |||||||||||||
Db  297 GGCTCGGCTTCCACACAGCTATATGCCCTCGTGCACACTCGCGCCGGCGAGC 238

Qy  557 TCGTGCTGCTGCACGTACCTCAAGGTGTGGGCGCCTTCCTGCTTCCATTGCAAGGCGGCA 616
        || ||| |||||||||| ||||||||| ||| ||| ||| |||||||| |||||||||||
Db  177 TCGTTGCTCTCGCAAAACGTACCTCAAGGTGGGCCCTTCCACTGCAGTTGCAAGGCGCA 118

Qy  617 TCGACGGTGAACCATGCACGACGCGTGGTGGTGCTGCTGCTGGTGGACCTGCACCCCAG 668
        || |||| || ||||||| | ||||| || |||||||| ||||| ||||| |||||| ||
Db  117 TCGACGTGATCACCATGCAGACGCGCTGTGTGCTGTTGGTGGACATACACCCCAG 66
```

Middle / right column

```
REFERENCE  2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
  PUBMED    10899143
REFERENCE  3  (bases 1 to 1005)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..1005
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="061L14"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG061DF07LP1-end : T7"
```

```
ORIGIN
Query Match    32.1%;  Score 325.4;  DB 9;  Length 1005;
Best Local Similarity   69.7%;  Pred. No. 5.5e-58;
Matches   469;  Conservative   1;  Mismatches  198;  Indels  5;  Gaps  2;

Qy    1 ATGAACTCGTGGGGACCGGGGCCTGGCGGGGGCGGGGGCTACTGGTGGGGCACGGATGGGGGTCTCGGCTG 60
        ||| | | ||| | | ||| |||||||||||||||| ||||||||||||||| |||||| |||||||||||
Db    6 AGGGACGCGGCGGGACTTAGTTGCGTCGCGTTGGTTTTGGGATTAHTTCGGTGTGTTCTG 65

Qy   61 CTGTGTCCAACG--CGCTGGTGTGTGCTCGTGCCTGCTGCCTCTGGTTGGTTTTCTCGGGGGGGTCGGCTG 119
        ||||||||||||  ||||||||| |||| ||| |||||||| |||||||||||||||||| |||||||||
Db   66 CTCTCCAACGTGCAGTTGTTCTGATCTGCTTTCTCGTGTACAACCCCGGAGATCCGGCAAACAGGT 125

Qy  120 GCCGCGGCGCTTCACCCTGAAACCTCACGTGCGGGAAACCTGCTGTTGCACCGTGGTCAACCGTCAACAT 179
        |||||||||| ||||| ||| ||||||||||| | |||| |||| |||||| ||||| |||||||||||
Db  126 ACCGGGGGCGTTTTCACTCTCAAACTTGACGTTTTGCAACCTGTTGCAGTGGCTGTCAACAT 185

Qy  180 GCCGCTCACGCTGGCCGCGGCGCGTTGCGTGGCGCAGCGGGCGACCCCGCTGTGCG 239
        |||| |||||||||||| || |||||||| ||| ||||||| | | ||||||||||
Db  186 GCCCCCTGACCCCTGGCGTCACCACGGGCGCAGCGGCAGCGGCCGGCTTCTGCCA 245

Qy  240 CCTCGCTGCCTTCCTCCGACACACCTTCCTGGCTGCGCCAACTCCATGCTCAGCATGGCCGCGG 299
        ||| |||||| |||||| | || ||| | ||| | || | ||||||||||||||||| ||| |
Db  246 GCTTGTGGGTTCCTCGACACCTTCCTCACCACCAACGCCATGCTGAGCATGGCAGCGGCT 305

Qy  300 CAGCATCGACCGGCTGGGTGGGGGTGGCGGTGGTGTTCCCCTCGGTGTTCCCCGCCAAGCCGCT 359
        ||||||||||||||||||||| || |||| |||||||| ||| | ||| |||| || |||||||
Db  306 CAGCATCGACCGGTGGGGGTGGCCGTGGTGTTCCCTCCGCGCTACCACTCCAGAATCCGGCA 365

Qy  360 CCGCGGCGGCGCTCATGGTTGGCCGTACACGTGGCGTCGCACGCGGCGCTTCACCTTCCCAGCGCG 419
        |||||||| | |||||| || || |||||| | || | ||| ||||||| |||||| ||||| |
Db  366 CCGGGACGGCGGATCGGCGTGGGGATACACGGCCGCTGCTGCTCCCCTTCTCCACGCGT 425

Qy  420 CGCGGCTCGCGCCCCTGTCTCGGCGTCGGGTTCCACCAGTGTAGGCCTCGTGCACGCGTGCAG 479
        ||||| ||| | |||| |||||| |||||| ||||| ||| | || || |||| |||||||||
Db  420 CGCGGGCCACCGTGCCTCCTGGGGTCGGCTTGGCGTCAACGTCGTGGCTGCTCCAGCTGCAG 479

Qy  480 CCGGGCGGCGACGAGCGGCGC----TGGCGTTCGCGCCGTCTTCACTCGGCGGCTTCCACGCTC 535
        ||||| ||||| | ||||       ||||||  || ||| |||||||||||| || |||| |||
Db  486 CGCAAGCGCGAAGGGAGACCAGGTTGGTTGGTTTGTGTTTTTCACGGCGGCTTTGACTCTC 545

Qy  536 TCAGCTTCCTGCTCTCCTTCGTCGTGCTGCTGCACGTACCTCAAGGTGCTCAAGGTGG 595
        | ||||||||| |||| |||||| |||||||||||| |||||||||||||| |||||||
Db  536 TCAGCTTCCTTGCTCTTCGGTCGTGCTGCTGCACGTACCTGAAAGTGGGTGCTCGTGGG 605

Qy  596 CCGGCGCTTCCATTGCAAGGCGCATCGACGGTGATCACCATGCACCATTGCACACGCGCTGTTGGTGG 655
        |||| ||| ||| || ||||||||| ||| || || ||| ||| ||||| | |||| ||||||||||
Db  546 TCACCTTCCTGCTCACGGCTGATCGGTTGCTGGTTGTGTTGTCTGCTGGAAGTGCTCAAAGTGG 605
```

RESULT 13

```
Db    606 CCAGATTTCACTGTAAACGATCAGCTGACGTATCACGATGACGCGCTGGTGCTGCTGGTGG 665
              |||||||
Qy    656 ACCTGCACCCCAG 668
              ||||||||||||
Db    666 ACATTCACCCCAG 678
```

LOCUS   T09060   318 bp   mRNA   linear   EST 03-AUG-1993
DEFINITION   EST06953 Infant Brain, Bento Soares Homo sapiens cDNA clone HIBB055
              5' end, mRNA sequence.
ACCESSION   T09060
VERSION   T09060.1 GI:390088
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 318)
  AUTHORS   Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.
  TITLE   Rapid cDNA sequencing (expressed sequence tags) from a
              directionally cloned human infant brain cDNA library
  JOURNAL   Nat. Genet. 4, 373-380 (1993)
  MEDLINE   94004965
  PUBMED   8401585
COMMENT   Contact: Adams, MD
              The Institute for Genomic Research
              932 Clopper Road, Gaithersburg, MD 20878
              Tel: 301869056
              Fax: 301869423
              Email: mdadams@tigr.org
              Seq primer: M13 Reverse.
FEATURES   Location/Qualifiers
     source   1..318
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="ATCC (inhost):85542"
                     /db_xref="taxon:9606"
                     /clone="HIBB055"
                     /clone_lib="Infant Brain, Bento Soares"
ORIGIN

Query Match   29.6%; Score 299.8; DB 7; Length 318;
Best Local Similarity   97.8%; Pred. No. 1.2e-52;
Matches 312; Conservative   0; Mismatches   6; Indels   1; Gaps   1;

Qy 640 CTGGTGCTGCTGGTGGGACCTGCACCCCAGTGTGCGGGGAACCGTGTCTGGAGGAGCAGAAG 699
            |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1 CTCGTGCTGCTGGTGGGACCTGCACCCCAGTGTGCGGGGAACGNTGTCTGGAGGAGTAGAAG 60

Qy 700 CGGAGGCGACAGCGAGCCACCAAGAAAGATCAGCACCTTCATPAGGACACTTCCTTGTGTGC 759
            |||||||||||||| ||||||||||||||||||||||||||| ||||||||||||||||||
Db   61 CGGAGGCGACAG-GAGCCACCAAGAAGATCAGCACCTTCATPAGGACACTTCCTTGTGTGC 119

Qy 760 TTCCGCGCCCTPATGTTGATCACCAGGCTGTGCCACGTGCCATCGCGTGCCATCGCGCTCC 819
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 120 TTCGGCGCCCTTATGTTGATCACCAGGCTCTTCTCCACGGTGCCATCGCGTGCCATCGCGCTCC 179

Qy 820 CACTGGGGGGTGCTGTCCAAGTGTCTTGGCGTAGCAGCAAGGCCGCATCCGACCCCTTTTGTG 879
            |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db 180 CACTGGGGGGTGCTGTCTCCAAGTGCTTNGCGTACAGCAAGGCCGCATCCGACCCCTTTTGTG 239

Qy 880 TACTCCTTACTGGACACCAGTACCGCCAAAAAGCTGCAAGGAGATTCTGAACAGGCTTCCTG 939
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 240 TACTCCTTACTGCGACACCAGTACCGCCAAAAAGCTGCAAGGAGATTCTNAACAGGCTTCCTG 299

Qy 940 CACACGACGCTCCCATCCACT 958
            |||||||||||||||||||||
Db 300 CACAGACGNTCCATCCACT 318
```

RESULT 14

LOCUS   CNS03YUE   1081 bp   DNA   linear   GSS 01-SEP-2000
DEFINITION   Tetraodon nigroviridis genome survey sequence T7 end of clone
              069M05 of library G from Tetraodon nigroviridis, genomic survey
              sequence.
ACCESSION   AL266783
VERSION   AL266783.1 GI:7988568
KEYWORDS   GSS; genome survey sequence.
SOURCE   Tetraodon nigroviridis
  ORGANISM   Tetraodon nigroviridis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
              Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
              Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
              Saurin,W. and Weissenbach,J.
  TITLE   Estimate of human gene number provided by genome-wide analysis
              using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
  PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
              Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
              Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE   Characterization and repeat analysis of the compact genome of the
              freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
  PUBMED   10899143
REFERENCE   3 (bases 1 to 1081)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT   This sequence is a single read and was generated as part of a large
              scale clone-end sequencing project of the Tetraodon nigroviridis
              genome. For more information, please take a look at
              http://www.genoscope.cns.fr/Tetraodon.
FEATURES   Location/Qualifiers
     source   1..1081
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /clone="069M05"
                     /clone_lib="G"
                     /db_xref="taxon:99883"
                     /note="Genoscope sequence ID : C0BG069AG03LP1-end : T7"
ORIGIN

Query Match   26.6%; Score 270; DB 9; Length 1081;
Best Local Similarity   63.4%; Pred. No. 2.4e-46;
Matches 411; Conservative   1; Mismatches   236; Indels   0; Gaps   0;

Qy   21 CCTGGCGCGGGGGGCTACTGGTGGGCACGATGGGGGTCTCGCTGGTCTGTCCAACGGCGGTGGTGCT 80
            |||||| |||||||||| ||||||| ||||||||||| |||||| |||||
Db 166 CCTGGAAGTGTGCGCCTCCTGTGACGACACCGCGGTGGTGGTCTCGGTGTTGACCAACTTGTCCGTGCT 225

Qy   81 GCTCTGCCTGCTGCACGAGCGCGGACAATCCGCGCGCAGGCGCCTCCTCTTCACCGTGAA 140
            ||| ||| |||||| || ||| |||| |||| ||||||||||||| |||||||
Db 226 GCTATGCTACTGCCAGAGCAGCCACGGAGTTAAGATCCCACGTTCCCGGCCATCTCATCCTCAA 285

Qy 141 CCTCACGTGCGGAGACCTTGCTGTGTGACCAGCGTGGTCAACATGCGCGTCACGCTGGGCGCCGGCGT 200
            ||||||||||||| ||||||| |||| ||||| |||| |||||||| |||| ||||||
Db 286 CCTCTCCTTCTCCAACATCCTCCTGCCGCCGCCATCAACACATCCGACGAGACGTTCCTGGCGAGT 345

Qy 201 CGTGGCGCCAGCGACGGCCAGCCGGGGGGGGG 260
            || ||||| | |||||||||||||
Db 346 GGCCGAAAAAACGGCAAGCCCTTCGGGGGACCCGTTCTGCCAAGACGGCCAGCTTTGCCGAGAC 405

Qy 261 CTTCCTGGCTCCCAACTCCATGCTCAGCATGGCATGGCGCGGCTCCAGCGCCGTGCAGCGCCCCTGCGCCTGGGTGGC 320
```

```
COMMENT      This sequence is a single read and was generated as part of a large
             scale clone-end sequencing project of the Tetraodon nigroviridis
             genome. For more information, please take a look at
             http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..870
```

```
REFERENCE    3  (bases 1 to 870)
  AUTHORS    Genoscope.
  TITLE      Direct Submission
  JOURNAL    Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
             BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
             - Web : www.genoscope.cns.fr)
REFERENCE    2
  AUTHORS    Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
             Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
             Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE      Characterization and repeat analysis of the compact genome of the
             freshwater pufferfish Tetraodon nigroviridis
  JOURNAL    Genome Res. 10 (7), 939-949 (2000)
  MEDLINE    20359837
   PUBMED    10899143
REFERENCE    1
  AUTHORS    Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
             Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
             Saurin,W. and Weissenbach,J.
  TITLE      Estimate of human gene number provided by genome-wide analysis
             using Tetraodon nigroviridis DNA sequence
  JOURNAL    Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE    20296633
   PUBMED    10835645
```

```
ACCESSION    AL182059
VERSION      AL182059.1  GI:7820150
KEYWORDS     GSS; genome survey sequence.
SOURCE       Tetraodon nigroviridis
  ORGANISM   Tetraodon nigroviridis
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
             Tetraodontoidea; Tetraodontidae; Tetraodon.
```

```
LOCUS       CNS025GY/c            870 bp    DNA     linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            236J14 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
```

```
RESULT 15
CNS025GY/c
```

```
Qy    766 CGTGATCACGGTGCAGACCCTGCTGCTGGTGGTGGATATTCACCCCAG 813
          |||||||||||||||||||||||||||||||||||||||||||||||
Db    766 CGTGATCACGGTGCAGACCCTGCTGCTGGTGGTGGATATTCACCCCAG 813
```

```
ORIGIN
        /organism="Tetraodon nigroviridis"
        /mol_type="genomic DNA"
        /db_xref="taxon:99883"
        /clone="236J14"
        /clone_lib="G"
        /note="Genoscope sequence ID : C0AG236DE07SP1-end :
         PUC-Ori"

Query Match     24.1%;  Score 244.4;  DB 9;  Length 870;
Best Local Similarity   62.5%;  Pred. No. 5.7e-41;
Matches  405;  Conservative  5;  Mismatches 236;  Indels  2;  Gaps  2;
```

```
Qy     21 CCTGGCGCGGGGGGGCTACTGGTGGGGCACGGATNGGGCGGGCCCGCGTCCGCGTCGTTC 80
          ||||: ...
Db    562 TTTTCTCACCACCAACGCAAATGCTTCGAGCATGGGCCC ...

Qy    261 CTTCCTGGGCTGCCAACTCCATNGCTTCANGATNGGGCG 320
Db    562 TTTTCTCACCACCAACGCAAATGCTTCGAGCATGGGCGG 503

Qy    621 GGCCGAAAACGGGCAAAGGCTTCGGG-GACCCGTTCTGCCAAGCGGGC 668
Db    621 GGCCGAAAACGGGCAAAGGCTTCGGG-GACCCGTTCTGCCAAGCGGGC 563
```

```
(rotated alignment columns — see original)
```

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Run on: July 5, 2005, 13:24:11 ; Search time 733 Seconds
         (without alignments)
         8676.664 Million cell updates/sec

Title: US-10-735-991-5
Perfect score: 1014
Sequence: 1 atgaactcgtgggacgcggg..........ttctgcggtgtctgagtga 1014
Scoring table: Gapop 10.0 , Gapext 1.0
Sequence: IDENTITY NUC
Searched: 6313374 seqs, 3116092125 residues
Total number of hits satisfying chosen parameters: 12626748
Maximum DB seq length: 200000000
Minimum DB seq length: 0
Post-processing: Minimum Match 0%
         Maximum Match 100%
         Listing first 45 summaries

Database :   Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*	1006.2	99.2	1728	21
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*	986.8	97.3	1014	14
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*	868.4	85.6	1172	19
4:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*	472.2	46.6	585	10
5:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*	418.2	41.2	1092	16
6:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*	418.2	41.2	1092	19
7:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*	418.2	41.2	1092	21
8:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*	418.2	41.2	1092	19
9:	/cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*	416.6	41.1	1352	13
10:	/cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*	416.6	41.1	1092	13
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*	416.6	41.1	1092	13
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*	416.6	41.1	1092	13
13:	/cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*	416.6	41.1	5665	14
14:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*	415	40.9	1092	16
15:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*	415	40.9	1092	16
16:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*	415	40.9	1092	21
17:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*	415	40.9	1092	13
18:	/cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*	415	40.9	1092	19
19:	/cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*	415	40.9	1092	19
20:	/cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*	415	40.9	1092	16
21:	/cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*	415	40.9	1092	16
22:	/cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*	415	40.9	1092	20
23:	/cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*	415	40.9	1092	19
24:	/cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*	415	40.9	1092	19
25:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*	415	40.9	1092	21
26:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*	415	40.9	1163	9
27:	/cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*	415	40.9	1188	10
28:		415	40.9	1188	21
29:		415	40.9	1188	21
30:		415	40.9	1510	13
31:		415	40.9	1510	14
32:		415	40.9	1510	14
33:		415	40.9	1510	14
34:		415	40.9	1510	14
35:		415	40.9	1510	14
36:		415	40.9	1510	14
37:		415	40.9	1510	14
38:		415	40.9	1510	14
39:		415	40.9	1510	14
40:		415	40.9	1510	14
41:		415	40.9	1510	14
42:		415	40.9	1510	14
43:		415	40.9	1510	14
44:		415	40.9	1510	14
45:		415	40.9	1510	14

ALIGNMENTS

Sequence 172, Appl	US-10-505-486-172	
Sequence 97, Appl	US-10-083-168-97	
Sequence 3, Appli	US-10-735-991-3	
Sequence 26254, A	US-09-918-995-26254	
Sequence 31, Appl	US-10-321-807-31	
Sequence 31, Appl	US-10-321-807-31	
Sequence 31, Appl	US-10-314-048A-31	
Sequence 9, Appli	US-10-080-960-9	
Sequence 7, Appli	US-10-080-960-7	
Sequence 1, Appli	US-10-166-221-1	
Sequence 2, Appli	US-10-296-121-2	
Sequence 6, Appli	US-10-088-726-6	
Sequence 95, Appl	US-10-321-807-95	
Sequence 95, Appl	US-10-321-807-95	
Sequence 95, Appl	US-10-314-048A-95	
Sequence 70, Appl	US-10-149-826-70	
Sequence 95, Appl	US-10-897-815-95	
Sequence 1, Appli	US-09-736-131-1	
Sequence 185, App	US-09-782-974C-185	
Sequence 185, App	US-10-467-492A-185	
Sequence 589, App	US-10-975-979-185	
Sequence 589, App	US-10-052-586-589	
Sequence 589, App	US-10-174-590-589	
Sequence 589, App	US-10-176-758-589	
Sequence 589, App	US-10-175-737-589	
Sequence 589, App	US-10-174-581-589	
Sequence 589, App	US-10-176-483-589	
Sequence 589, App	US-10-176-749-589	
Sequence 589, App	US-10-176-914-589	
Sequence 589, App	US-10-176-915-589	
Sequence 589, App	US-10-173-706-589	
Sequence 589, App	US-10-175-738-589	
Sequence 589, App	US-10-175-752-589	
Sequence 589, App	US-10-176-482-589	
Sequence 589, App	US-10-176-757-589	
Sequence 589, App	US-10-176-913-589	
Sequence 589, App	US-10-180-552-589	

RESULT 1
US-10-735-991-5

; Sequence 5, Application US/10735991
; Publication No. US20040121395A1
; GENERAL INFORMATION:
; APPLICANT: GOODNOW, Jr., Robert Alan
; APPLICANT: MARK, David Fu-Chi
; APPLICANT: MARTIN, Mitchell Lee
; APPLICANT: ROSINSKI, James Andrew
; TITLE OF INVENTION: Sequence #115 as a target for identifying weight modulating
; TITLE OF INVENTION: compounds
; FILE REFERENCE: 21366
; CURRENT APPLICATION NUMBER: US/10/735,991
; CURRENT FILING DATE: 2003-12-15
; PRIOR APPLICATION NUMBER: US 60/436,375
; PRIOR FILING DATE: 2002-12-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.2
US-10-735-991-5

; SEQ ID NO 5
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: GPCR #115
; LOCATION: (1)..(1014)
; OTHER INFORMATION: AJ505757

Query Match   100.0%;   Score 1014; DB 19; Length 1014;
Best Local Similarity   100.0%;   Pred. No. 2.6e-253;

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	1014	100.0	1014	19	US-10-735-991-5	Sequence 5, Appli
2	1011	99.7	1011	19	US-10-735-991-1	Sequence 1, Appli
3	1009.2	99.5	2530	19	US-10-757-262-129	Sequence 129, App
4	1009.2	99.5	2581	9	US-09-742-732-1	Sequence 1, Appli
5	1009.2	99.5	2581	21	US-10-742-732-1	Sequence 1, Appli
6	1006.2	99.2	1011	9	US-09-742-732-3	Sequence 3, Appli
7	1006.2	99.2	1011	21	US-10-921-613-3	Sequence 3, Appli

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

RESULT 2
US-10-735-991-1

; Sequence 1, Application US/10735991
; Publication No. US20040121395A1
; GENERAL INFORMATION:
; APPLICANT: GOODNOW, Jr., Robert Alan
; APPLICANT: MARK, David Fu-Chi
; APPLICANT: MARTIN, Mitchell Lee
; APPLICANT: ROSINSKI, James Andrew
; TITLE OF INVENTION: Sequence #115 as a target for identifying weight modulating
; TITLE OF INVENTION: compounds
; FILE REFERENCE: 21366
; CURRENT APPLICATION NUMBER: US/10/735,991
; CURRENT FILING DATE: 2003-12-15
; PRIOR APPLICATION NUMBER: US 60/436,375
; PRIOR FILING DATE: 2002-12-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.2
;
; SEQ ID NO 1
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: GPCR #115
; LOCATION: (1)..(1011)
; OTHER INFORMATION: CHR7-36867
US-10-735-991-1

Query Match   99.7%; Score 1011; DB 19; Length 1011;
Best Local Similarity 100.0%; Pred. No 1.6e-252;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1014; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

[DNA sequence alignment blocks — Qy (query) and Db (database) rows in paired format, positions 1–1014, not individually legible]

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (186)...(1199)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2530)
; OTHER INFORMATION: n = A,T,C or G
US-10-757-262-129

Query Match   99.5%;   Score 1009.2;   DB 19;   Length 2530;
Best Local Similarity   99.7%;   Pred. No. 4.9e-252;
Matches 1011;   Conservative   0;   Mismatches   3;   Indels   0;   Gaps   0;

```
Qy    1  ATGAACTCGTGGGACACGGGGCCGGGGCCTGGGAGGCGGGCCGGGGGGGGGGGGGGCTCGGTG      60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  186  ATGAACTCGTGGGACACGGGGCCTGGGAGGCGGGCCGGGGGGGGGGGCTACTGGTGGGGGCTC     245

Qy   61  CTGTCCAAGCGGCTGGTTGCTCCTCGCCCTGAGCTACCGGGCCCAAGATGCCCCTC     120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  246  CTGTCCAACGGGCTGGTTGCTCGTCTCGCCCTGCCACGCGCGGGACCATCCGCGGCG     305

Qy  121  CCGGCGCTTCTTCACCCTGAACCTCAGTGCGGGAACCTGTCGTCGCGTCACATG     180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  306  CCGGGGCTCTTCACCCCTGAACCTCACGTGCGGGAACCTGCTCTGCACCGTGGTCAACATG     365

Qy  181  CCGGCTCACGCTGGCGCGCGGCGTCGGTGCGCAGCGGGCAGCGCGACCGGCCTGTGCGCG     240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  366  CCGCTCACGCTGGCCGTCGGGCGTGGGCGCGGCAGCCGGCCGCAGCGGCCTGTGCCGCG     425

Qy  241  CTGGCTGCCTTCCTCGACACACCTTCCTGGCTGCGGCGGGGCCTGGGCGGGGGGGGGGGCTC     300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  426  CTGGCTGCCTTCCTCGACACACCTTCCTGGCTGCGCCAACTCCATGCTCAGCCATGCAGCCC     485

Qy  301  AGCATCGACACCCTGGGTGGCCGGTGGTCTTCCCGCTGAGCGTACCGGCGCCAAGATGCCCTC     360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  486  AGCATCGACACCCCGCTGGGGTTGGGCCGTGGGTCCCGGCGGCCAAGATGCGGCCTC     545

Qy  361  CGGCGACGGCGGCGGCTCAGTTGGTGGGCCTTACACGGGCTGGCCAGCGGCTCACCTTCCCGCC     420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  546  CGGCGACGGCGGCGGCTCATGGTGGCGCCTTACACGGGGCTCACCGTCACCTTCACGCGCGC     605

Qy  421  GCGCTCGCCCCTGTCGTCTCGGCGGCTTCCACCAGGTCACCGGCGCGTCGTGTGTCAGC     480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  546  GCGCTCGCCCTGTCGTCTCGGCGGCTTCCACCAGGTCACCGGCGCGTCGTGGTCAGC     605

Qy  481  CGGCGCCAGACGAGCGCGCGCTGCGCCGCGTTCGCCGTCTTCACTGCGCCGCTTCCCAGC     540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  666  CGGCGGCCGGACGAGCGCGCTCGCCGCTTCGCCGTCTTCCACTGCGCCGCTCTCAGC     725

Qy  541  TTCCTGCTCTTCCTTCGTGTGGCTCTGCTGCACGGTACCTCAAGGGTGGCCCGC     600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  726  TTCCTGCTCTTCCTTCGTGTCGTGGCTCTGCTGCACGGTGCTCAAGGTGGGCCCGC     785

Qy  601  GCGCTCGACCCCTGTCGTCGGCTCGGCGTTCACCAGAGTCGTACCGGCTCGTGTGCAGC     660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  606  GCGCTCGCCCTGTCCGTGGCTCGGCGTTCCACCAGAGTCGTACCGGCTCGTGTGCAGC     665

Qy  661  CACCCCAGTGTGTCGGGGAACCGCTGTCTGGGAGGAGCAGAAGCCGAGGCGGACGAGCCACC     720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  666  CACCCCAGTGTGTCGGGGAACCGCTGTCTGGGAGGAGCAGAAGCCGAGGCGGACGAGCCACC     725

Qy  721  AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTCTTCGCGCCGCCTATCGTGATCACC     780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  AAGAAGATCAGCACCTTCATAGGGACCTTCCTTGTGTCTTCGCGCCGCCTATGTGATCACC     780

Qy  781  AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTGCCCACTGGGGGGTTGCGTGTCCAAG     840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTGCCCACTGGGGGGGTGCGTGTCCAAG     840

Qy  841  TGCTTGGCGGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGCGACACCAG     900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  TGCTTGGCGGTACAGCAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGCGACACCAG     900

Qy  901  TACCGCAAAAGCTGCAAGGAGATTCTGAACAGGCTCTCCATGCTCCATCCACTCC     960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TACCCCAAAAGCTGCAAGGAGATTCTGAACAGGCTCTCCATGCTCCATCCACTCC     960

Qy  961  TCTGGGCCTCACAGGGGGACTCTCACAGCCAGAACATTCTGCCGGGTGTCTGAG    1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  TCTGGCCTCACAGGGGGACTCTCACAGCCAGAACATTCTGCCGGGTGTCTGAG    1011
```

RESULT 3
US-10-757-262-129
; Sequence 129, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheti, Venkateswarlu
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
; TITLE OF INVENTION: 3751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; CURRENT APPLICATION NUMBER: US/10/757,262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440,318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444,783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457,901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468,775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471,614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478,742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488,529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491,156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499,594
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506,332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 129
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens

; FILE REFERENCE: MP103-007P1RNOMNIM
; FEATURE:
; NAME/KEY:
; LOCATION:
; FEATURE:
; NAME/KEY:
; LOCATION:

```
Qy    901  TACCGCAAAAAGCTCCAAGGAGATTCTGAACAGGCTTCCTGCGACACAGAGACGCTTCCATCCACTCC  960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1086  TACCGCAAAAAGCTGCAAGGAGATTCTGAACAGGCTTCCTGCGACAGAGACGGCTTCCATCCACTCC 1145

Qy    961  TCTCGGCCTCACAGGCGACTCTCACAGCCAGGAACATTCTGCCGGTGTCTCGAGTGA 1014
                |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   1146  TCTCGGCCTCACAGGCGACTCTCACAGCCAGGAACATTCTGCCGGTGTCTCGAGTGA 1199
```

RESULT 4

US-09-742-732-1

Sequence 1, Application US/09742732

Patent No. US2002005137A1

GENERAL INFORMATION:

APPLICANT: M. Alexandra Glucksmann and Keith Robison

TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/742,732

FILING DATE: 20-Dec-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/175,983

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Attorney, Jean M. Silveri

REGISTRATION NUMBER: 39,030

REFERENCE/DOCKET NUMBER: MNI-036

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2581 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 184..1194

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-742-732-1

Query Match   99.5%; Score 1009.2; DB 9; Length 2581;

Best Local Similarity   99.7%; Pred. No. 4.9e-252;

Matches 1011; Conservative   0; Mismatches   3; Indels   0; Gaps   0;

```
Qy      1  ATGAACTCGTGGAGGACGGCGGCGCGCTGGCCGGGCCTACTGGTGGGCCACCGCTCTCGGCTG   60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184  ATGAACTCGTGGAGGACGGCGGCGGGCCCTGGGCGGGGCTACTGGTGGGCACCGGCTCTCGGCTG  243

Qy     61  CTGTCCAACGGCGGCTCGGTGTTCGCTTGCTCCTGCTCGCACAGCGGGCACATCCCGCCCAGGCG  120
                |||||||||||||||||||| ||| |||||||||||||||||||||||||||||||||||||
Db    244  CTGTCCAACGGCGGCTCGGTGCTCGCCTCTGCTCGCACAGCGGGACATCCCCGCCCAGGCG  303

Qy    121  CCGGGCGCTCTTCAACCCTGAACCTCAGCGTGCGGGAGAACCTGCTGTGTCACCGTGGTCAACATG  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304  CCGGGCGCTCTTCACCCTGAACCTCAGGTGCGGGAACCTGCTGTGTCGACCGTGGTCAACATG  363

Qy    181  CCGCTCACGTCTGGCCGGCGTCGTGGCGGCAGCGGAACCGCGGGGGGGGGGGGGCAGCG  240
```

RESULT 5

US-10-921-613-1

Sequence 1, Application US/10921613

Publication No. US20050079550A1

GENERAL INFORMATION:

APPLICANT: M. Alexandra Glucksmann and Keith Robison

TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 28 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```
Db    364  CCGGCTCACGGCTGGCCGGCGTCGTGGCCGGCCGGGGCGGCGGGCGCAGCCGACCGGCTGTTGCGC  423

Qy    241  CTGGGCTGCCTTCCCTCGGACACCTTCCTGCGCCGCGCGCCAACTCCAATGCTCAGCAATGCTCAGCA  300
                || ||||||| ||||| |||| ||||||||||||||||||||||||||||||||||||||||
Db    424  CTGGGCTGCCTTCCTCGGACACCTTCCTGGGCTGCCAACTCCAATGCTCAGCATGGCCCGCGCTC  483

Qy    301  AGCATCGACCCGCTGGGTGGGCCGTGGGCCGTGTCTTCCCGGCTGAGCTTACCGGGGCCAAGATGCGCCTC  360
                ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484  AGCATCGACCCGCTGGGTGGGCCGTGGGCCGTGGTCTTCCCGGCTGAGCTTACCGGGGCCAAGATGCGCCTC  543

Qy    361  CGGCAGCGGCCGCGCTCATGGTTGGCCTTACAGCGTGGCTGCACGCGGCTCACCTTCCCAGCGCC  420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544  CGGCAGCGGCCGCGCTCATGGTGGCCTACACGTGGCTGCACGCGGCTCACCTTCCCAGCGCC  603

Qy    421  GCGCTCGCCCTGGTCCTGGGGCTTCACCAGCAGCTCGTTGCACGCGGCTCGTGCCAGC  480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604  GCGCTCGCCCTGTCCCTGGGGCTTCACCAGCAGCTCGTTGCACGCGGCTCGTGCCAGC  663

Qy    481  CGGCGGCCGCCAGACGACGGCCCTGCGCCTTCAGCGTCTTCACTGGCGCTCTTCCAGCTCGCCAGC  540
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664  CGGCGGCCGCCGGACGAGGCCTGCGCTTCGCCTTCAGCTCTTCACTGGCGCTCTTCAGC  723

Qy    541  TTCCTGCTCTCCTTCGTTCGTGCTCTTGCTGCACGGTACCTTCAAGGTTGCTTCAAGGTGCCCGGC  600
                ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    724  TTCCTGCTCTCCTTCGTTCGTGCTCTTGCTGCACGTACCTTCAAGGTGCTTCAAGGTGCCCGC  783

Qy    601  TTCCATTGCAAGGCGCATCGAGCGTGATCACCATCAGAGAACCAGCGCCTCGTCGTCGTGGCTG  660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    784  TTCCATTGCAAGGCATCGAGCGTGATCACCATGCAGAGAACGCGCTGCGTCGTCGTGGCTG  843

Qy    661  CACCCCAGTGTGCGGGAACCGCTGTCTGCTGGAGGACGCAGAAGCCGAGGGGGACACGGAGCCACC  720
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844  CACCCCAGTGTGCGGGAACGGCTGTCTGCTGGAGGAGCAGAAAGCCGAGGGGGACACGGAGCGACC  903

Qy    721  AAGGAAGATCAGCACCTTCATTAGGGACTTCCTTGTGTTCGCGCCCCTGTATGTGATCATCACC  780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904  AAGAAGAATCACCACCTTCATTAGGGACTTCCTTGTTCGGCCCTGTATGTGATCACC  963

Qy    781  AGGGCTAGTGGAGCTCTTCTCCACGGTGCCCACTGGGGGTGCTTCCAAG  840
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db    964  AGGGCTAGTGGAGCTCTTCTCCACGGTGCCCACTGGGGGGTGCTTCCAAG 1023

Qy    841  TGCTTGGCGTACAGCAAAGGCCGCCATCGACCCCTTTGTTGTTGTACTCCTTACTGGACACCAG  900
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024  TGCTTGGCGTACAGCAAAGGCCGCCATCCGACCCCTTTGTGTACTCCTTACTGGACACCAG 1083

Qy    901  TACCGGCAAAAGCTGCAAGGAGGATTCTGAACAGGCTCCTGCGACAGACGCTTCATCCACTCC  960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1084  TACCGCAAAAGCTGCAAGGAGAATTCTGAACAGGCTCCTGCACAGACGCTTCCATCCACTCC 1143

Qy    961  TCTCGGCCTCACAGGCGACTCTCACAGCCAGGAACATTCTGCCGGTGTCTGAGTGA 1014
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144  TCTCGGCCTCACAGGCGACTCTCACAGCCAGGAACATTCTGCCGGTGTCTGAGTGA 1197
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/10/921,613
  FILING DATE: 18-Aug-2004
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/09/013,634
  FILING DATE: 26-JAN-1998
ATTORNEY/AGENT INFORMATION:
  NAME: Attorney, Jean M. Silveri
  REGISTRATION NUMBER: 39,030
  REFERENCE/DOCKET NUMBER: MNI-036
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (617)227-7400
  TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2581 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 184..1194
  SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-921-613-1
```

Query Match　　　99.5%;　Score 1009.2;　DB 21;　Length 2581;
Best Local Similarity　99.7%;　Pred. No. 4.9e-252;
Matches 1011;　Conservative　0;　Mismatches　3;　Indels　0;　Gaps　0;

```
Qy      1 ATGAACTCGTGGGACGCGCGGCCTGGCGGGGGCTACGTGGTGGGCACCATGGGCGTCTCGGCTG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    184 ATGAACTCGTGGGACGCGCGGCCTGGCGGGGGCTACGTGGTGGGCACCATGGGCGTCTCGGCTG  243

Qy     61 CTGTCCAACGGCGCTGGGTGCTGCTGCCTGCTCGACAGGCGCATCGCGCGCCAGGCGCCACC  120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    244 CTGTCCAACGGCGCTGGGTGCTGCTGCCTGCTCGACAGGCGCATCGCGCGCCAGGCGCCACC  303

Qy    121 CCGGCGCTCTTCACCCTGGAACCTCACGTGGGGGAACCTGCTGGGTCAACATG  180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    304 CCGGCGCTCTTCACCCTGGAACCTCACGTGGGGGAACCTGCTGGGTCAACATG  363

Qy    181 CCCGCTCACCGTGGGCGGCCTCGTGGGCGCAGCCGCGGACCGGCTCTTGTGCCGC  240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    364 CCCGGTCACCGTGGGCGGCCTCGTGGGCGCAGCCGCGGACCGGCTCTTGTGCCGC  423

Qy    241 CTGGCTGCCTTCCTCCGACACCTTCCCTGGCGTGCCAACTCCATGCTCGAGCTC  300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    424 CTGGCTGCCTTCCTCCGACACCTTCCCTGGCGTGCCAACTCCATGCTCGAGCTC  483

Qy    301 AGCATCGACCGCGGCTGGGGTGGCCGGCTGGGCCGGGGCTGAGCTGCGCGCTC  360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    484 AGCATCGACCGCGGCTGGGGTGGCCGGCTGGGCCGGGGCTGAGCTGCGCGCTC  543

Qy    361 CGCCACGCGCGGCTCATGTGGCCTACACGTGGCCTCACGCTCACCTTCCCAGCCGCC  420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    544 CGCCACGCGCGGCTCATGTGGCCTACACGTGGCCTCACGCTCACCTTCCCAGCCGCC  603

Qy    421 CGGCTCGCCCTGTCCTGGCTGGCTTCCACCAGCTGTACGCCCTCGTCGCACGGCGTGTGCAGC  480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    604 CGGCTCGCCCTGTCCTGGCTGGCTTCCACCAGCTGTACGCCCTCGTCGCACGGCGTGTGCAGC  663

Qy    481 CGGCGGCCAGACGAGCGCCGCTGGCGTCGCGTTCGCGCGTCTGCGCGTCTCAGG  540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    664 CGGCGGCCGGACGAGCGCCGCTGGCGTCGCGTTCGCGCGTCTGCGCGTCTCAGG  723

Qy    541 TTCCTGCTCTCCTTCGTCGTGCTCTGCTGCACGTACCTCAAGGTTGCTCAAGGTGGCCCCGC  600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    724 TTCCTGCTCTCCTTCGTCGTGCTCTGCTGCACGTACCTCAAGGTTGCTCAAGGTGGCCCCGC  783
```

US-09-742-732-3
Sequence 3, Application US/09742732
Patent No. US20020055137A1
GENERAL INFORMATION:
  APPLICANT: M. Alexandra Glucksmann and Keith Robison
  TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR
  NUMBER OF SEQUENCES: 14
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: LAHIVE & COCKFIELD, LLP
    STREET: 28 State Street
    CITY: Boston
    STATE: Massachusetts
    COUNTRY: USA
    ZIP: 02109
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/742,732
    FILING DATE: 20-Dec-2000
    CLASSIFICATION: <Unknown>
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: 09/175,983
    FILING DATE: <Unknown>
  ATTORNEY/AGENT INFORMATION:
    NAME: Attorney, Jean M. Silveri
    REGISTRATION NUMBER: 39,030
    REFERENCE/DOCKET NUMBER: MNI-036
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (617)227-7400
    TELEFAX: (617)742-4214
  INFORMATION FOR SEQ ID NO: 3:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 1011 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: cDNA
    FEATURE:
      NAME/KEY: CDS

```
Qy    601 TTCCATTGCAAGCGCATCGACCGTGATCACCATGCAGCAGACGCGCTGGTGGTGCTGGTTGGGACCTG  660
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    784 TTCCATTGCAAGCGCATCGACGTGATCACCATGCAGCAGACGCGCTCGTGCTGGTGCTGGTTGGGACCTG  843

Qy    661 CACCCCCAGTGTGCGGGGAACGCTGTCTCTGGAGANGCAGANGCGGAGGGGCGGACAGGCGGCCACC  720
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 CACCCCCAGTGTGCGGGGAACGCTGTCTCTGGAGAGACAGAAGGCGGAGGGGCGGACAGGCGGCCACC  903

Qy    721 AAGAAGATCAGGCACCTTCATAGGGACCTTTCCTTTGTTGTGTTGGCTTCGCGCGGCCCTATGTGATCACC  780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 AAGAAGATCAGGCACCACCTTCATAGGGACCTTTCCTTTGTTGTGTTGGCTTCGCGCGGCCCTATGTGATCACC  963

Qy    781 AGGCTAGTAGTGGAGCTCTTCTCCACGCGTGCCCATCGGCTCCCCACTCGTGGGCGGGTGGTGCTGTTCCAAG  840
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db    964 AGGCTAGTGGAGCTCTTCTCCACGCGTGCCCATCGGCTCCCCACTCGGGGGGTTGCTGTTCCAAG  1023

Qy    901 TACCGCAAAAGCTGCAAAGGAGGGAGATTCTGAACAGGCTCCTTGCACAGAGAGGCTCCATTCCACTCC  960
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1084 TACCGCAAAAGCTGCAAAGGAGGGAGATTCTGAACAGGCTCCTTGCACAGAGAGGCTCCATTCCACTCC  1143

Qy    961 TCTGGGCCTCACAGGCGACTCTCACGACCCAGAACACATTCTGCGGGTGGTGTCTGAGTGA  1014
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 TCTGGGCCTCACAGGCGACTCTCACGACCCAGAACACATTCTGCGGGTGGTGTCTGAGTGA  1197
```

```
;
US-09-742-732-3
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
    LOCATION: 1..1011

Query Match    99.2%;    Score 1006.2;  DB 9;  Length 1011;
Best Local Similarity
Matches 1008;  Conservative    0;  Pred. No. 2.8e-251;  Mismatches   3;  Indels    0;  Gaps    0;

Qy      1 ATGAACTCGTGGGACAGCGGGGGCCTGGGCGGGGGCTACTGGTGGGCACGGATGGGGTTCGGCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAACTCGTGGGACAGCGGGGGCCTGGGCGGGGGCTACTGGTGGGCACGGATGGGGTTCGGCTG 60

Qy     61 CTGTCCAACGCGCCTGGTTGCTCCTGCTGCTGCGGGGGACATCCGCCGCCCAGGGCG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGTCCAACGCGCCTGGTTGCTCCTGCTGCTGCGGGGGACATCCGCCGCCCAGGGCG 120

Qy    121 CCGGCGGCTCTTCACCCTGAACCTCAGTGTCGCGGGAACCTCAGCCCTGGTCAACATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCGGCGGCTCTTCACCCTGAACCTCAGTGTCGCGGGAACCTCAGCCCTGGTCAACATG 180

Qy    181 CCGGCTCACGCTGGCGCGGTGTGGGCCAGCGACGGCGACCGCCTGTGCCGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCGGCTCACGCTGGCGCGGTGTGGGCCAGCGACGGCGACCGCCTGTGCCGC 240

Qy    241 CTGGCTGCCTTCCTCGACACCTTCCTGTGCGGCGCGGCGGGCCAGCGCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTGGCTGCCTTCCTCGACACCTTCCTGTGCGGCGCGGCGGGCCAGCGCTC 300

Qy    301 AGCATCGACGCCGTGGGGCCGTGGTCTTCCCGCTGAGCTTACCGGGCCAAGAGATCGCGCCTC 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGCATCGACGCCGTGGGGCCGTGGTCTTCCCGCTGAGCTTACCGGGCCAAGAGATCGCGCCTC 360

Qy    361 CGCGACGCGGGCGCTCATGGTGGCCTACACCGTGGCCTGCACCGCCGC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CGCGACGCGGGCGCTCATGGTGGCCTACACCGTGGCCTGCACCGCCGC 420

Qy    421 GCGCTCGCCCCTGTCCTGGCTCGGGCTTCCACCAGCTGTACGCCTGTGCAGC 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCGCTCGCCCCTGTCCTGGCTCGGGCTTCCACCAGCTGTACGCCTGTGCAGC 480

Qy    481 CGGCGGGCCAGCGACGACGCCTGCCTTCGCCGTCTTCACTGGCTCTCAGC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CGGCGGGCCAGCGACGACGCCTGCCTTCGCCGTCTTCACTGGCTCTCAGC 540

Qy    541 TTCCTGCTCTCCTTCGTCGTGCTCTGCGCAGTACCTCAAGGTGGCCCCGC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCCTGCTCTCCTTCGTCGTGCTCTGCGCAGTACCTCAAGGTGGCCCCGC 600

Qy    601 TTCCATTGCAAGCGCATCGAGCTGATCAACCATGCAGACGCTCGTGCTGGTGGTGGACCTG 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TTCCATTGCAAGCGCATCGAGCTGATCAACCATGCAGACGCTCGTGCTGGTGGTGGACCTG 660

Qy    661 CACCCCAGTGTGCGGGAACCGCTGTTCTTGGAGGAGAGGCGACAGCGACGAGCCACC 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CACCCCAGTGTGCGGGAACCGCTGTTCTTGGAGGAGAGGCGACAGCGACGAGCCACC 720

Qy    721 AAGAAGATCAGCACACTTCATAGGGACACTTCCTTGTTGTTGCTTCGGCGCCTAGTGATCACC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AAGAAGATCAGCACACTTCATAGGGACACTTCCTTGTTGTTGCTTCGGCGCCTAGTGATCACC 780

Qy    781 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGGCGGGGGTGCTGTGTCCAAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGGCGGGGGTGCTGTGTCCAAG 840

Qy    841 TGCTTGGCGGTTACAGAGGAGGCCCATCGGACACCCCTTGTACTCCTTACCTGGGACACCAG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCTTGGCGGTTACAGAGGAGGCCCATCGGACACCCCTTGTACTCCTTACCTGGGACACCAG 900

Qy    901 TACCGGCAAAAGCTGCAAGGAGGAATTCTGAACAGGCTCCTCGACAGAGACGCTCCATCCACTCC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TACCGGCAAAAGCTGCAAGGAGGAATTCTGAACAGGCTCCTCGACAGAGACGCTCCATCCACTCC 960

Qy    961 TCTGGCCTCACAGGGCGACTCTCACAGCCAGAACAATTCTGCGGGTGTCTGAG 1011
```

```
Db    961 TCTGGCCTCACAGGGCGACTCTCACAGCCAGAACAATTCTGCGGGTGTCTGAG 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 7
US-10-921-613-3
; Sequence 3, Application US/10921613
; Publication No. US20050079550A1
; GENERAL INFORMATION:
;   APPLICANT: M. Alexandra Glucksmann and Keith Robison
;   TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR
;   NUMBER OF SEQUENCES: 14
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: LAHIVE & COCKFIELD, LLP
;     STREET: 28 State Street
;     CITY: Boston
;     STATE: Massachusetts
;     COUNTRY: USA
;     ZIP: 02109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/10/921,613
;     FILING DATE: 18-Aug-2004
;     CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/09/013,634
;     FILING DATE: 26-JAN-1998
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Attorney, Jean M. Silveri
;     REGISTRATION NUMBER: 39,030
;     REFERENCE/DOCKET NUMBER: MNI-036
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617)227-7400
;     TELEFAX: (617)742-4214
;   INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1011 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..1011
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-921-613-3

Query Match    99.2%;    Score 1006.2;  DB 21;  Length 1011;
Best Local Similarity
Matches 1008;  Conservative    0;  Pred. No. 2.8e-251;  Mismatches   3;  Indels    0;  Gaps    0;

Qy      1 ATGAACTCGTGGGACAGCGGGGGCCTGGGCGGGGGCTACTGGTGGGCACGGATGGGGTTCGGCTG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAACTCGTGGGACAGCGGGGGCCTGGGCGGGGGCTACTGGTGGGCACGGATGGGGTTCGGCTG 60

Qy     61 CTGTCCAACGCGCCTGGTTGCTCCTGCTGCTGCGGGGGACATCCGCCGCCCAGGGCG 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CTGTCCAACGCGCCTGGTTGCTCCTGCTGCTGCGGGGGACATCCGCCGCCCAGGGCG 120

Qy    121 CCGGCGGCTCTTCACCCTGAACCTCAGTGTCGCGGGAACCTCAGCCCTGGTCAACATG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CCGGCGGCTCTTCACCCTGAACCTCAGTGTCGCGGGAACCTCAGCCCTGGTCAACATG 180

Qy    181 CCGGCTCACGCTGGCGCGGTGTGGGCCAGCGACGGCGACCGCCTGTGCCGC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCGGCTCACGCTGGCGCGGTGTGGGCCAGCGACGGCGACCGCCTGTGCCGC 240

Qy    241 CTGGCTGCCTTCCTCGACACCTTCCTGTGCGGCGCGGCGGGCCAGCGCTC 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Query Match        99.2%;  Score 1006.2;  DB 21;  Length 1728;
Best Local Similarity   99.7%;  Pred. No. 2.9e-251;
Matches 1008;  Conservative 0;  Mismatches 3;  Indels 0;  Gaps 0;

Qy    1  ATGAACTTCGTGGGAACGGGGGGGGGCCTGGGGGGGGCGGGGGCTTACTCGGTGGTGCGCACGGATNGGGGTGCTCGGCTG   60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    1  ATGAACTTCGTGGGAACGGGGGGGGGCCTGGGGGGGGCGGGGGCTTACTCGGTGGTGCGCACGGATGGGGGTGCTCGGCTG   60

Qy   61  CTGTCCAACGCGGCGGTGGTTGCTGCTCTGCCTGTGTGCACAGCGGCGGGGAC               120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  CTGTCCAACGCGGCGGTGGTTGCTGCTCTGCCTGTGTGCACAGCGGCGGGGAC               120

Qy  121  CCGGCGCGTCTTCACCGTCAGGTGCGCGGGGAACCTGCTGGTCACCGTTGGTCAACATG        180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  CCGGGCGTCTTCACCGTCAGGTGCGCGGGGAACTTGCTGGTGCACCGTGGTCAACATG        180

Qy  181  CCGCTCACGCTGGCCCGGTCGTGGGCCGGAGGGGCAGGGGCAGGCGGTGCGC         240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CCGCTCACGCTGGCCCGGTCGTGGGCCGGAGGGGCAGGGGCAGGCGGTGCGC         240

Qy  241  CTGGCTGCCTTCCTCGACACTTCCATGCTCAGCATGGCCCGCGCCTC              300
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  CTGGCTGCCTTCCTCGACACTTCCATGCTCAGCATGGCCCGCGCCTC              300

Qy  301  AGCATCGACACCGCTGGGTGGTCTTCCCGGCTCGGCCACCGGGCCAGC         360
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  301  AGCATCGACACCGCTGGGTGGTCTTCCCGGCTCGGCCACCGGGCCAGC         360

Qy  361  CGGGCGGCCAGACGGAGCGGCGCCTGCCGCTTCGCCGTCACCAGGCGGCCGC       420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361  CGGGCGGCCAGACGGAGCGGCGCCTGCCGCTTCGCCGTCACCAGGCGGCCGC       420

Qy  421  GCGCTCGCCCTGTGTCCTGGCCTGCCTTCGCACGCTCGTGGCAGAGTGCTGCAGC      480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421  GCGCTCGCCCTGTGTCCTGGCCTGCCTTCGCACGCTCGTGGCAGAGTGCTGCAGC      480

Qy  481  CGGGCGGCCCGGACGAGCGCCTCGGCGTTCGGCCGTTCCACAGCGGCGGCCTC       540
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481  CGGGCGGCCCGGACGAGCGCCTCGGCGTTCGGCCGTTCCACAGCGGCGGCCTC       540

Qy  541  TTCCCGTCTCCTTCGTTCGTGTCCTGCGCAGTACCTCAAGGTGGCTCAAGGTGGGCCGC      600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541  TTCCCGTCTCCTTCGTTCGTGTCCTGCGCAGTACCTCAAGGTGGCTCAAGGTGGGCCGC      600

Qy  601  TTCCATTGCAAGGCAATCGACGTGATCACCATGCAGACGCTCGTGTTCGCCCTAGTGTGATCACC  660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601  TTCCATTGCAAGGCAATCGACGTGATCACCATGCAGACGCTCGTGTTCGCCCTAGTGTGATCACC  660

Qy  661  CACCCCAGTGTGCGGGAAACGCTGTCTGGAGGAGCGGAAGCGGAGGGGACAGCACC    720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  661  CACCCCAGTGTGCGGGAAACGCTGTCTGGAGGAGCGGAAGCGGAGGGGACAGCACC    720

Qy  721  AAGAAGATCAGCACCTTCATAGGGACTTCCTTGTTGCTTGCGCGGTGATGATCACC         780
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  721  AAGAAGATCAGCACCTTCATAGGGACTTCCTTGTTGCTTGCGCGGTGATGATCACC         780

Qy  781  AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTCCCACTGGGTGCTCCAAG         840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  781  AGGCTAGTGGAGCTCTTCTCCACGGTGCCCATCGGCTCCCACTGGGTGCTCCAAG         840

Qy  841  TGCTTGGCGTACAGCAAGGCCGCCATCGGACCCCTTTGTGTACTCCTTACTCGGACACCAG       900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  TGCTTGGCGTACAGCAAGGCCGCCATCGGACCCCTTTGTGTACTCCTTACTCGGACACCAG       900

Qy  901  TACCGGCAAAAGCTGCAAAGGAGATTCTGAACAGGCGCTCCATCCACTCC           960
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TACCGGCAAAAGCTGCAAAGGAGATTCTGAACAGGCGCTCCATCCACTCC           960

Qy  961  TCTGGGCCTCACAGGGCGACTCTCACAGCCCAGAACATTCTGCCGGTGTGTCTGAG   1011
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  TCTGGGCCTCACAGGGCGACTCTCACAGCCCAGAACATTCTGCCGGTGTGTCTGAG   1011
```

RESULT 9
US-10-083-168-97
; Sequence 97, Application US/10083168
; Publication No. US20030023069A1
; GENERAL INFORMATION:
; APPLICANT: Liiaw, Chen W.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Behan, Dominic P.
; APPLICANT: Maciejewski-Lenior, Dominique
; APPLICANT: Leonard, James N.
; APPLICANT: Ortuno, Daniel
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Endogenous And No. US20030023069A1-Endogenous, Constitutively Act
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0320
; CURRENT APPLICATION NUMBER: US/10/083,168
; CURRENT FILING DATE: 2002-02-26
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 97
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-083-168-97

Query Match 97.3%; Score 986.8; DB 14; Length 1014;
Best Local Similarity 98.3%; Pred. No. 3e-246;
Matches 997; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy   1 ATGAACTCGTGGGGACGGCGGCCGCTACTGGTGGCCACCATGGGCGGTCTCGCTG   60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 ATGAACTCGTGGGGACCGGCGGCCGCTACTGGTGGCCACCATGGGCGGTCTCGCTG   60

Qy   61 CTGTCCAACGCGCTGGTGTGCTGTGCTTCTGCACGCGCGGGACATCCGCGGCCAGGCG   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 CTGTCCAACGCGCTGGTGTGCTGTGCTTCTGCACGCGCGGGACATCCGCGGCCAGGCG   120

Qy   121 CCGGGCGGCTCTTCACCTGAACCTCAGGTGCGGGGAACCTGGTGCATGGTCAACATG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 CCGGGCGGCTCTTCACCTGAACCTCAGGTGCGGGGAACCTGGTGCATGGTCAACATG   180

Qy   181 CCGCTCACGTGGCCGGCCGGTCGTGCCGCAGGCGCACCGGCCCTGTGCCCGC   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCGCTCACGTGGCCGGCCGGTCGTGCCGCAGGCGCACCGGCCCTGTGCCCGC   240

Qy   241 CTGGCCTGGCCTTCCTCGACACACTTCCTGGCCTGCCAACTCCATGCTCAGCAT   300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 CTGGCTGCCTTCCTCGACACACTTCCTGGCTGCCAACTCCATGCTCAGCAT   300

Qy   301 AGCATCGACACCGCTGGGGTGGGCCGTGGTCTTCCCGCTGAGCTACCGGCCGCCTC   360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCATCGACACCGCTGGGGTGGGCCGTGGTCTTCCCGCTGAGCTACCGGCCGCCCT   360

Qy   361 CCGAGATGCGCGCCTCATCCTGGCCTACACGTGTCACCTTCCCAGCCCGCC   420
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CCGGCGAGCGCGCCTCATCCTGGCCTACACGTGTCACCTTCCCAGCCCGCC   420

Qy   421 GCGCTCGCCCTGTCTGGCCTCGGCTTCCACCAGCTCCTCGTTGCACCTGTGC   480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCGCTCGCCCTGTCTGGCCTCGGCTTCCACCAGCTCCTCGTTGCACCTGTGC   480

Qy   481 CGGCCGCCAGACGAGCGCCCTGCCCTTCGCCGTCTTCACGCCTCTCAGC   540
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CGGCCGCCGGACGAGCGCCTGCCCTTCGCCGTATTCACTGGCCTCTCAGC   540

Qy   541 TTCCTCGCTCTCTCGTTCGTGTTGTCCGTGATCACCAC   600
          ||||||||||||||||||||||||||||||||||||||
Db   541 TTCCTCGCTCTCTCCTTCGTGTTGTCCGTGATCACCAC   600

Qy   601 TTCCATTGCAAGGCGCATCGAGACGCTCGTGCTGCTGCTTCTGGTGGACCTG   660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 TTCCATTGCAAGGCGCATCGAGACGCTCGTGCTGCTGCTTCTGGTGGACCTG   660

Qy   661 CACCCCCAGTGTGCCGGGAACCGCTGTCGTCTGGAGGAGGAGGAAGGCGGAGCCACC   720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CACCCCCAGTGTGCCGGGAACCGCTGTCGTCTGGAGGAGGAGGAAGGCGGAGCCACC   720

Qy   721 AAGAAGGATCAGCACCTTCATAGGGAGACACCTTCCTTGTGTCTGCTTCGGCGGCCCT   780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AAGAAGGATCAGCACCTTCATAGGGAGACACCTTCCTTGTGTCTGCTTCGGCGGCCCT   780

Qy   781 AGGCTAGTGGAGCTCTTCTCCACGGTGTCCCATCGGCCTCCCACTGGGGGGGTGTCCAAG   840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGGCTAGTGGAGCTCTTCTCCACGGTGTCCCATCGGCCTCCCACTGGGGGGGTGTCCAAG   840

Qy   841 TGCTTGGCGTGCAGCAAGGCCGCATCCGACCCCTTGTGTACTCCTTACTGGCAGACACCG   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGCTTGGCGTGCAGCAAGGCCGCATCCGACCCCTTTGTGTACTCCTTACTGGCAGACACCAG   900

Qy   901 TACCCGCAAAAGCTGCAAAGGAGAATTCGACAGGCTCCTGCCACAGAAGCCTCCATCCTCC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TACCCGCAAAAGCTGCAAAGGAGAATTCGACAGGCTCCTGCCACAGAAGCCTCCATCCACTCC   960

Qy   961 TCTGGCCTCACAGGCGACTCTCACAGCCCAGAACATTCTGCCGGTGTCTGAGTGA   1014
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 TCTGGCCTCACAGGCGACTCTCACAGCCCAGAACATTCTGCCGGTGTCTGAGTGA   1014

RESULT 10
US-10-735-991-3
; Sequence 3, Application US/10735991
; Publication No. US20040121395A1
; GENERAL INFORMATION:
; APPLICANT: GOODNOW, Jr., Robert Alan
; APPLICANT: MARK, David Fu-Chi
; APPLICANT: MARTIN, Mitchell Lee
; APPLICANT: ROSINSKI, James Andrew
; TITLE OF INVENTION: Sequence #115 as a target for identifying weight modulating
; TITLE OF INVENTION: compounds
; FILE REFERENCE: 21366
; CURRENT APPLICATION NUMBER: US/10/735,991
; CURRENT FILING DATE: 2003-12-15
; PRIOR APPLICATION NUMBER: US 60/436,375
; PRIOR FILING DATE: 2002-12-23
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 3
; LENGTH: 1172
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: GPCR #115
; LOCATION: (1)..(1172)
; OTHER INFORMATION: Q90XI3
US-10-735-991-3

Query Match 85.6%; Score 868.4; DB 19; Length 1172;
Best Local Similarity 91.0%; Pred. No. 1.6e-215;
Matches 923; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

Qy   1 ATGAACTCGTGGGACGGCGGGGCCTGGCGGGGGGCTCTGCTG   60
          ||||||||||||||||||||||||||||||||||||||||||
Db   27 ATGAACTCGTGGGACACGGGGCCTGGCGGGGGGCTCTGCTG   86

Qy   61 CTGTCCAACGCGCTGGTGTGCTGCTGCCTGCCGGGCGGGGACAATCGCGCCAGGCG   120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   87 CTGTCCAACGCGGCTGGTGGTCTGCTGCCTCCGCACAGGGCCACAATCGCGCCAGGCG   146

Qy   121 CCGGGCGGCTCTTCACCTGAACCTCAGGTGCGGGGAACCTGTTGCACGGTCAACATG   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   147 CCGGGGCGGCTCTTCACCTCAAACCTCACGTGTCGGGCGCAACCTGCGTGGTCAACATG   206

Qy   181 CCGCTCACGTCGGCCGCGGTCGTGGCCGCAGGCGCGGGCCGCGCCTGTGTCCGC   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   207 CCACTAACACTGGCCCGCGTCGTGGCACAAGCGGACCGCGCGCCTTGTCCGCGC   266

Qy   241 CTGGCTGCCTTCCTCCGCACACCTTCCTGGCCTGCCTGCCAACTCCCATGGCCCGGCGGCTC 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   267 CTGGCTGCCTTCCTCCGACACCTTTCTGGCCCGCCCCAACTCCCATGGCTATGGCCCGGCGCTC 326

Qy   301 AGCAGTCGACCGCGCTGGGTGTGGCCGTGGTCTTCCCGGCCCTGAGCTTACCGGCGCCAAGCATCCTC 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   327 AGCAGTCGACCGCGCTGGGTGGTGTTGGTCTTCCCGGCTACCGGTCCCAAGATGCCCTC 386

Qy   361 CGGCGACGCGGGCGGCGGCTCATGGGTGGCCTACACGTGGCTGGCAGCCGCCCTC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   387 CGAGATGCGCGCCCTTCATGGTGGTGTTCCTGGCTACACGTGCCAAGATGCCCACC 446

Qy   421 GCGCTCGCCCCTGTCCTGGCCTTCCACAGCTGTCAGCCGCTGTGCCAGC 480
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   447 GCGCTCGCCCCTGTCCTGGCCTTCCACAGCTGTATGCCCGGCTGGCAGC 506

Qy   481 CGGCGCGGCCCAGCAGCGAGCGGCGGCGCTTGCGGCTTCGCCGGCCTTCCACGGCTCTCAGC 540
           ||||||||||||||||||||||||||||||||||||||||||
Db   507 CGGGCGGCCCGACGAGGCCCTGCGCTTTGCTGTCTTCCACCAGCGCCATCGGCGCTTAGC 566

Qy   541 TTCCTGCTCTCCTTGGTGTGTGCTGCTGCGACGTACCTCAAGGTGCTCAAGGTGGCCCGC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   567 TTCCTGCTCTCCTTCATGGTGCTTCTTGTCTTCACCAGCGCTCAAGGTGCTCAAGGTGGCCCGGT 626

Qy   601 TTCCATTGCAAGCGCATCGACGTGATCACCATGACGAGACGCTGCTCGTCGGTGACCTG 660
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   627 TTCCACTGCAAGCGGCATCGACGTGATCACCATGACGAGAGACGGTCGTCGTTAGTGACATC 686

Qy   661 CACCCCAGTGTGCGGGGAACCGTGTCTGGAGGAGCGAGAAGCGAGACGAGCGCCACC 720
           ||||||||||||||||||||||||||||||||||||||||||||
Db   687 CATCCCAGTGTGAGGGAACGATGTGTGAGGAGCAGAAGCGTGCTCGGAGGAGCAGAAGCGAGAGGC 746

Qy   721 AAGAAGATCAGCACTTCATAGGGACCTTCCTTGTTGTTGTGTGTGATCACC 780
           ||||||||||||||||||||||||||||||||||||||||||
Db   747 AAGAAGATCAGCACACTTCATAGGGACCTTCCTCGTGTGTCTTTGCACCTATGTGATTACC 806

Qy   781 AGGCTAGTGGAGCTCTTCTCTCCACGGTGCCCATCGGGCGGGGTGCCCAAG 840
           ||||||||||||||||||||||||||||||||||||||||||
Db   807 AGGTGGTGGGAACTCTTCTCCACAGCCCATTGACTGACTCGTGGGTGTGTCCAAG 866

Qy   841 TGCTTGGCGGGCTACAGCGAAGGCCGCATCCGACCCCTTTGGTGGTACTCCTTACTGGACACCAG 900
           ||||||||||||||||||||||||||||||||||||||||
Db   867 TGCTTGGCCTTACAGCAAGGCTGCTTCTGACCCCTTGCTGGTGTACTCCTTGCTGCGACCCAG 926

Qy   901 TACCGCAAAGCTGCAAGGAGGATTCTGAACAGGCTCCTGCACAGACGCCTCCATCCACTCC 960
           ||||||||||||||||||||||||||||||||||||
Db   927 TACCCGCAGGAGCTGCAAGGAGCTTCTGAACAGGATCTTCAAACAGACGGCTCCATTCACTCT 986

Qy   961 TCTGGCCTCACAGGCGGACTCTCACAGCCAGAAACAATTCTCGCCGGGTGTCTCGAGTGA 1014
           |||||||||||||||||||||||||||||||||||||
Db   987 GTGGGCCCTCACAGTGACTCTCACAGCCAGAAACATTCTGCCAGTGTCGCGAATGA 1040

Query Match 46.6%; Score 472.2; DB 10; Length 585;
Best Local Similarity 99.4%; Pred. No. 1e-112;
Matches 474; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy   527 TCCACCGCTCTCAGCTTCTCCTCGTGCTTCCTCGTGCTGCTACGTACCTCAAGGTGC 586
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   49 TCCACCGCTCTCAGCTTCCTGCTTCCTCGTGCTTCGTGCTGCTACGTACCTCAAGGTGC 108

Qy   587 TCAAGGTGGCCCGCGCTTCATTGCAAGCGGCATCGACGTGATCACCATGACGAGACGCTGC 646
           ||||||||||||||||||||||||||||||||||||||||||||
Db   109 TCAAGGTGGCCCGCGCTTCCATTGCAAGCGGCATCGACGTGATCACCATGACGAGACGCTGC 168

Qy   647 TGCTGTGGGACCTGCACCCCAGTGTGCGGGAAACGCTGTCTGGAGGAGCAGAAGC 706
           ||||||||||||||||||||||||||||||||||||||||||||
Db   169 TGCTGGTGGGACCTGCACCCCAGTGTGCGGGGAACGCTGTCTGGAGGAGCAGAAGC 228

Qy   707 GACAGCGGAGCCACCAAGAAGAGATCAGCACCTTCATAGGGACCTTCCTTCGTTCGCTTCGCGC 766
           ||||||||||||||||||||||||||||||||||||||||||||
Db   229 GACACGGAGCCACCAAGAAGAGATCAGCACCTTCATAGGGACCTTCCTTCGTTCGCTTCGCGC 288

Qy   767 CCTATGTGATCACCAGGCTAGTGGAGCTCTTCTCTCCACGGTGCCCATCGGGCGGG 826
           ||||||||||||||||||||||||||||||||||||||||
Db   289 CCTATGTGATCACCAGGCTAGTGGAGCTCTTCTCTCCACGGTGCCCATCGGGCGGG 348

Qy   827 GGGTGCTTGGCGGGCTACAGCAAGGCCGCGCCATCCGACCCCTTTGTGTGGTACTCCT 886
           ||||||||||||||||||||||||||||||||||||||||||
Db   349 GGGTGCTTGGCCGGGCTACAGCAAGGCTGCTTCTGACCCCTTGCTGGTGTACTCCT 408

Qy   887 TACTCGGACACCAGTACCCGCAAAAGCTGCAAGGAGGATTCTGAACAGGCTCCTGCACAGAC 946
           ||||||||||||||||||||||||||||||||||||||||||
Db   409 TACTCGGACACCAGTACCCGCAAAAGCTGCAAGGAGGATTCTGAACAGGATCTTCAAACAGAC 468

Qy   947 GCTCCATCCACTCCTCTGGCCTCACAGGCGGGACTCTCACAGCCAGAAACAATTCTCGCCGG 1003
           ||||||||||||||||||||||||||||||||||||||||||||
Db   469 GCTCCATCCACTCCTCTGGCCCTCACAGGCGGGACTCTCACAGCCAGAAACATTCTGCCGG 525

```
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/181,749
; PRIOR FILING DATE: 2000-02-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 31
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-321-807-31

Query Match             41.2%;  Score 418.2;  DB 16;  Length 1092;
Best Local Similarity   64.5%;  Pred. No. 1.1e-98;
Matches  624;  Conservative  0;  Mismatches  343;  Indels  0;  Gaps  0;

Qy    1 ATGAACTCGTGTGGGACCGCGGCGGCCCTGGCGGGCTACTGGTGGGGGCGTCTCGCTG  60
          |||||                  ||| ||||||||    ||||   |||    ||||||||
Db    1 ATGGGCCCCGACGGAGGCGCTGCGTCCTGGCGGGGTTCTCCCTCGGTGATGGGTACTGGGCCGTGGCGCTG  60

Qy   61 CTGTCCAACCGCGGCGTGGTGTGCGTGCTGCCTGCGCTGCACAGCGCCGACATCCCGCCCAGGCG  120
          |||||       ||||||   |||||  |||   |||||  |  |||||     ||||  ||||
Db   61 CTATCCAACCGCACTGGTGCTGCCTTTGTTGCCCTACCAGCGCTGAGCTCCGCACTCGAGCC  120

Qy  121 CCGGCGCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGGTGCACCGTTGGTCAACATG  180
          |||||||||| ||| |||||||||||   |||  ||||||   |||||||||   |||||
Db  121 TCAGGCGTCTCTCCTCCGGTGAATTCGTGGTGCCCACCTTGCTGCTGGGGCTGGACATG  180

Qy  181 CCGCTCACGCTGGCCGCGGTCGTGTGGGGCGCCAGCGGGGGGGACCCCCTGTGCCGC  240
          |||||||  ||||||||||    ||||||||||||||||   ||||||||||||||||
Db  181 CCCTTCACGCTGCTCCGGTGTGAATGGGGGGGACACCGTTCGGGGGCCCCCCATGCCAA  240

Qy  241 CTGGCTGCCTTCCTCCGGACACCTTCTGGCGTGCCAACTCATGCTCAGCATGGCGCGCTC  300
          ||||  |||  |||||||  |||    |||||||||   |||   ||||||||||||||
Db  241 GTCATTGGCTTCCTGGACACCTTCCTGGGGCTCCAACGGCGGTGGGCGGGCTG  300

Qy  301 AGCATCGACCGCTGGTGGTGGCCGTGGTCTTCCCGGCTACCGGGCCCTC  360
          |||  |||||||||||   |||||||   ||||||||||   |||||||
Db  301 AGCCGCAGACCAGTGGCGCGGCTTCCCCACTGCGCTATCGGCCCGACCG  360

Qy  361 CGCGACGGCGGCGGCTCATGGTGGCCTACACCGTGGCGCTCACGCTGCCGCC  420
          |||||| |||||   |||||||   ||||||||||||   |||||||||||
Db  361 CGCTATGCCGGCCCTGGCTGGGGGGACAGTCGCGTCGGCCTTCTCAGGCGCT  420

Qy  421 GCGCTCGCGCCCTGTCTCGGGCTTTCCACCAGCTGTACCCTCGTGTGCAGCGCTGTGCAGC  480
          |||||   ||||   |||||||  |||||||||   |||||   |||||    ||||||
Db  421 CCGCCCGAGCCTGAGCCGGTTCGCAGCGGCCCTTCGCCCGTCCAGTGGCTGTGGC  480

Qy  481 CGGCGGCCAGACGAGGAGCGCTTCGCCGCTTCTCTTCACCGCTCTCAGC  540
          |||||||||||||||||   ||||||||   |||||||||||   ||||
Db  481 CCGCCCGAGCCTGACGGTCGCGGCGTTCGCGCTGTTCGGCCGTGGGC  540

Qy  541 TTCGTGCCTCGGGGCATCGAGCCCTACACCATGCAGGCGCGGGCGTGGCCCGC  600
          ||||   |||||||||   ||  ||||||| ||||||  ||||||||||||
Db  541 TTCGGTCCGCCGGGGTCGCTCTGGCGCCTTCCAGGTTGCACGCGGGTGGGCACGG  600

Qy  601 TTCCATTGCAAGCGCATCGACCGTGATCACCATGCAGGACGCGGTGCTGCTGGACCTG  660
          |||| ||   ||||||||   ||||||||   |||   |||||||||||||||||||
Db  601 AGCCACTGCCAGCGCATGGACACCGGTGATCACCATGAAGGGCGCTGCGCTGCCGACTG  660

Qy  661 CACCCCCAGTGTGCGGGGAACGGCTGTCTGGACGGCAGAGAAGGGCGAGCCACC  720
          |||||     |||||||    ||||  |||||||| |||||||| ||| ||||||
Db  661 CACCCCAGTGTGCGGGAACGGCTGTCTGGACGGAGGAGCAAGAGCGGCGACGAGGCCAC  720

Qy  721 AAGAGAACACAGCAACACTTCATAGGGACCTTCCTGGTGTTCGCGCCCTATGTGTCACCAC  780
          |  | ||  ||||   |||||||   ||||||  ||||||||||||  |||||||||||
Db  721 AGGAAGATTGGCAATTGCCTATGCACACCTTTCTGCCCGGGATGTCATGAGGACC  780

Qy  781 AGGCTTAGTTGGAGCTCTTCTTCACGCGTGCCCATCGGGGGGTTGCTGTCTCCAAG  840
          ||||     ||||| |||  |||||   ||||      ||||||   |||  ||||
Db  781 AGGCTCGGCGGAGCTCGTGCCCTTCGTCACCCTGAAGCGCCAGTGGGGCATCCTCAGCAAG  840

Qy  841 TGCTTGGCGTACAGCAAGGCCCGATCCGACCCCTTTGTGTACTCCTTACTGCGACACCAG  900
          ||||||||||||||||||  |||||||   ||||||||||| |||||||||| ||||||
Db  841 TGCCTGACCTACAGCAAGGCTGCCGGTGCCGGTTCACGTTACTCTGCTTCCGCCGGCCG  900
```

```
Db  841 TGCCTGACCTACAGCAAGGCTGCCGGTGCCGGTTCACGTTACTCTGCTTCCGCCGGCCG  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  901 TACCCGCAAAGCTGCAAGGAGGATTCTGAACAGCGCCGTTCAGCGTTCCATCCACTCC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901 TTCCGCCCAAGTCCTGGCCGGCCATCGTGCACGGGTGCTCGAAGAGAACCCCGGCCCCAGA  960

Qy  961 TCTGGCC  967
           |||||||
Db  961 TCCACCC  967
```

```
RESULT 13
US-10-321-807-31
; Sequence 31, Application US/10321807
; Publication No. US20040137563A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruopong
; APPLICANT: Dang, Huong T.
; APPLICANT: Lowitz, Kevin P.
; TITLE OF INVENTION: Non-Endogenous, Constitutively Activated Human G Protein-Coupled
;   Receptors
; FILE REFERENCE: AREN0086
; CURRENT APPLICATION NUMBER: US/10/321,807
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: US/09/714,008
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,088
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,099
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,369
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/171,902
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/171,901
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/171,900
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/181,749
; PRIOR FILING DATE: 2000-02-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 31
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-321-807-31

Query Match             41.2%;  Score 418.2;  DB 19;  Length 1092;
Best Local Similarity   64.5%;  Pred. No. 1.1e-98;
Matches  624;  Conservative  0;  Mismatches  343;  Indels  0;  Gaps  0;

Qy    1 ATGAACTCGTGTGGGACCGCGGCGGCCCTGGCGGGCTACTGGTGGGGGCGTCTCGCTG  60
          |||||                  ||| ||||||||    ||||   |||    ||||||||
Db    1 ATGGGCCCCGACGGAGGCGCTGCGTCCTGGCGGGGTTCTCCCTCGGTGATGGGTACTGGGCCGTGGCGCTG  60

Qy   61 CTGTCCAACCGCGGCGTGGTGTGCGTGCTGCCTGCGCTGCACAGCGCCGACATCCCGCCCAGGCG  120
          |||||       ||||||   |||||  |||   |||||  |  |||||     ||||  ||||
Db   61 CTATCCAACCGCACTGGTGCTGCCTTTGTTGCCCTACCAGCGCTGAGCTCCGCACTCGAGCC  120

Qy  121 CCGGCGCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGGTGCACCGTTGGTCAACATG  180
          |||||||||| ||| |||||||||||   |||  ||||||   |||||||||   |||||
Db  121 TCAGGCGTCTCTCCTCCGGTGAATTCGTGGTGCCCACCTTGCTGCTGGGGCTGGACATG  180

Qy  181 CCGCTCACGCTGGCCGCGGTCGTGTGGGGCGCCAGCGGGGGGGACCCCCTGTGCCGC  240
          |||||||  ||||||||||    ||||||||||||||||   ||||||||||||||||
Db  181 CCCTTCACGCTGCTCCGGTGTGAATGGGGGGGACACCGTTCGGGGGCCCCCCATGCCAA  240
```

; TITLE OF INVENTION: for the Treatment of Metabolic-Related Disorders
; FILE REFERENCE: 22.US6.CIP
; CURRENT APPLICATION NUMBER: US/10/314,048A
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/096,511
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/995,543
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/399,917
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/404,761
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 60/410,747
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 31
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-314-048A-31

Query Match   41.2%; Score 418.2; DB 19; Length 1092;
Best Local Similarity   64.5%; Pred. No. 1.1e-98;
Matches 624; Conservative   0; Mismatches 343; Indels   0; Gaps   0;

Qy   1 ATGAACTCGTGGGACGCGGGCCGCCTGGGCGGGGGGCTACTGGGTGGGCACGANTGGGGGCTCTCGCTG 60
            |||| | ||||| |||| || | | ||| || | | || ||||| || ||||| |||||||
Db   1 ATGGGCCCCGGCGGAGGGCGGGGCTGCTGGGGCCGNTTCCTTGGTGATNGGGTACTGGGGCGTGGCGC 60

Qy   61 CTGTCAAACCGCGCTGGTGCTGTCCCTGCTGCACAGCGCGGGACAATCCCCGCCCAGGCG 120
            |||| ||||| |||||| | |||| |||| ||||| || ||||| | |||| |
Db   61 CTATCCAACCGACTCGGTCGTGCTTGGTTGTTGCCCCTACAGCGGCCTCCGGCGACTCGAGCC 120

Qy   121 CCCGGCGCTCTTCACCCTGAACCTCACGTGCGGGGAACCTGCTGTGCAGTCAACATG 180
            | || || |||||| | | ||| ||| ||| | | || ||| ||| ||| |||
Db   121 TCAGCGCGTCCTCCTGGTGAATCTGCGCGTGGCGCACCTGCTTGCTGCTGGGACATG 180

Qy   181 CCGCTCACGCTGGCCGCCGGCGTCGTGGCGCAGCGAGCCGCACCCGCCGTGTTGCCGC 240
            |||| || || ||||| ||||| |||||||| |||| || | ||| |||||| | ||
Db   181 CCCTTCACGGCTGCTGCTGTCGTGGATGCGGCACACCCGGGCCCCGCCGGCGCCAA 240

Qy   241 GTCATTGGCTTCCTGGACACCTTTCCTGGGGTCCAACGGGGCTCCGGGGGCCCGCTG 300
            || ||||| | |||||||||||| | || | || |||| | | || |||| |||
Db   241 GCCGAGAACCAGCAGTGCAACCGTGCGGCTTCCCACTCGCGGGGCCTGGGACCCGCTG 300

Qy   301 AGCAATCGACCGGTCGGGCTGGCGGTGGGCCCGGGCTACCGGGCGGCCCGGCCTC 360
            | | | |||| ||| ||| || ||| ||| |||| | | || | |||||
Db   301 AGCGAGAACCAGTGGCTGGGCCGTGGCGTCCCACTGGGCTTCCCGGCTCCGGCCTG 360

Qy   361 CGCGACGGCGGCGTCATGGTGGCCTACACGGTGGCCGGCGCTCACCTTCCCAGCCGCC 420
            | | ||| | || |||||| | |||| || ||||| | |||| || ||||| | ||
Db   361 CTATGCCGGACTCGTCCGGGTGGCCGTGGTGGCGCCGTCCCAGGTTGCACCGGGCAGC 420

Qy   421 GCGCTCGCGCCCTGTTCCTGGGCTTCCACCAGCGTGTACCGGCTGTGCAGC 480
            || || | | | | |||| | ||||||||||| | ||||| | |||
Db   421 GCACTTGGCTGCTCGTCGGCGTTGGTTGGCTTACAGCAGCCGGCTCCTGCCTG 480

Qy   481 CCGGCGCCCAGACGAGCGAGCGCCTTGCGCTTCGCGCGCTCTTCGCGCTCTCAGC 540
            | ||| | ||| | | | || | || | | | | | | ||| | | |
Db   481 CCGCCCCAGGCCTGAGGCGGTCCGGTGGCTTGGTTGCTACAGCAGCCTTCACGGTGC 540

Qy   541 TTCCTGCTCTCCTTCGTGTCGTCGCACGTACCTCAAGGTGCTCAAGGTGCCCCGC 600
            || || | ||||| | || || | ||||||||| |||||| |||| ||||||| |
Db   541 TTCGTGGCTGGCCGCGTGGGTGTTCCGTGGTACGCTCAAGGTTGCACGGGTTGCACGC 600

Qy   601 TTCCATTGCAAGCGCATCGATCACCAGAGCGCTTCGGCGGGTTGCTCGGTGGTGACCTG 660
            || | || | ||| | |||||| | | ||||| || | | | | | ||| |
Db   601 AGCCACTGCCAGCGCCATGGACACCGTGCGCTTCCGGCTTCCAGGTTGCACCCGGCAGC 660

Qy   661 CACCCCAGTGTGCGGGGAAGCTGTCTGGAGGACGAAGCGGAGCAGCGACGAGCCACC 720
            |||||||| ||| ||| || | | ||| || | || | ||| | | |||||
Db   661 CACCCCAGTGTGCGGGGACGCTGTCTGGAGGAGCAGGAGCAGCGCGGGCCGGGCCACC 720

```
Qy    721 AAGAAGATCAGCACCTTCATAGGGACTTCCTTGTTGGTCTTCGGCCGCCCTATGTGATCACC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AAGAAGATTGGCATTGCTATTGCGACCTTCCTCATCCTGCTTTTGTTCGGCCGCCCTATGTGATCACC 780

Qy    781 AGGCTAGTTGGGAGGCTCTTCTCCACGGTGCCCATCGGCGCTCCCCACTGGGGGGTGTCCTGTGTCCAAG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AGGCTGGGCGGAGCTCGTGCCCTTCGTCTCACCGTGAACGCCCCAGTGGGGCATCCTCAGCAAG 840

Qy    841 TGCTTGGCGTACAGCAAGAGGCCGCATCCGACCCCTTGTGTTGTACTCCTTACTGCGACACCAG 900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TGCCTGACCTACAGCAAGAGGGGTGGGGCCGGTTCACGTACTCTTCTGCTGCCGTTCGGCCG 900

Qy    901 TTCCGCCAAAAAGCTGCAAGGAGAGATTCTGACACAGACGCTTCATCCACTCC 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TACCGCAAAAGCTCGAAGGAGATTGTGCACAGACGACCCGGCCGCAGCA 960

Qy    961 TCTGGGCC 967
          ||  ||
Db    961 TCCACCC 967
```

RESULT 15
US-10-897-815-31

; APPLICANT: Unett, David J.
; APPLICANT: Chen, Ruoping
; APPLICANT: Richman, Jeremy
; APPLICANT: Connolly, Daniel
; APPLICANT: Dang, Huong T.
; APPLICANT: Choi, Bryan
; APPLICANT: Leonard, James
; APPLICANT: Hakak, Yaron
; APPLICANT: Liaw, Chen
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Lerner, Michael
; TITLE OF INVENTION: Human G Protein-Coupled Receptors and Modulators Thereof
; TITLE OF INVENTION: for the Treatment of Metabolic-Related Disorders
; FILE REFERENCE: 22.US6.CIP
; CURRENT APPLICATION NUMBER: US/10/897,815
; CURRENT FILING DATE: 2004-07-23
; PRIOR APPLICATION NUMBER: US/10/314,048
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/096,511
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/995,543
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 60/399,917
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/404,761
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 60/410,747
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 31
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-897-815-31

Query Match   41.2%;   Score 418.2;   DB 21;   Length 1092;
Best Local Similarity   64.5%;   Pred. No. 1.1e-98;
Matches   624;   Conservative   0;   Mismatches 343;   Indels   0;   Gaps   0;

```
Qy     1 ATGAACTCGTCGGGACCCGGGGCCTGGGCGGGGGCGCTACTGGTGGGGGGTCTCGGCTG 60
         ||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db     1 ATGGGGCCCCGGCGGGAGGGGGCTGGGCGGGGGTTCTTCCTTGGGTGATGGTACTGGGGCCGTGGGGCTG 60
```

```
Qy    61 CTGTCCCAACGCGGCTTGGTCGCTCTGCCTTGGGTGCTGCTGTGCTTCTGCGACAGCGGGCCGGACATCCGCGCCCGCCAGGCCG 120
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 CTATCCAAGGCACTGGGTGCTGCTGTGCTTTTGTTCGGCCCGCTACAGGCCGGGCCACCTTCGGCACTCGAGCC 120

Qy   121 CCGGCGGCTCTTCACCCTGACCTCACGTGGCGGGGAGACCTGCTGGTGGGGGGTCAACATG 180
         ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TCAGGCGGTCCTTCCCCGGTGAATCGTCGTCTAGGGCCCACCTTGCTCGGCGCACATG 180

Qy   181 CCGCTCACGCTGCGGCGGCCGCTCGTGTGCGCGACGAGCGGCCGGGCCGGACGCGTGCCGC 240
         ||    |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCCTTCACGTGCTCGGTTCGGTGTGCGGGGGCGCGACCGTGGCCGGGCCCGCCAA 240

Qy   241 CTGGCTGCCTTCCTTCCTGGACACCTTCCTGCTGTTGCCCAACTCCATGCTCAGCGATCC 300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 GTCATTGGCTTCCTGGACACCTTCCTGGCTGTCGCCAACGCGCCGTGGGGGGCTG 300

Qy   301 AGCCATCGACCGCTCGGTGGGCCGTGGTCTTCCCGGCTGAGCTACGGCGGGCCAAGGGCGGTGGACCTG 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGCCGCAGACCACCAGTGCGCTGGGGCAGTGGGGCCTTCCCCACTGCGCTACGGCGCGGC 360

Qy   361 CGCGACGCGCGCGCTCAGTGGTTGGGCCTACACGTGGGGTCGCAGGCGGCCGCGGCCAC 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 CGCTARGCGCCGGCCTGCTGCTGGGCGTTGCCGCGGGGACAGTCGCGTCGCCGGACGGC 420

Qy   421 GCGCGCTGCCCCTGTCCTGGCGCTTCCACCAGCTGTACCGGCTCGCACGGCTGGCGCAGC 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 GCACTTGGGCTCGCTGCTCGTGTGGGCTTGGCCTACAGCAGCGCCTTCGCGTTCGCGTTGCGGGC 480

Qy   481 CGGCGGCCAGACGAGCGCCTTCGCCGCTTCGGTCTCTTCACTGGCTCGCCTTCACGCTCTCAGC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 CCGCCGGAGCCTGAGGCGTCCGGCCTTCGCAGCCTTCACCCGGCACCTCCATGCCCTGGGC 540

Qy   541 TTCCCTGCTCTCCTTCGTTCGTGGCTCTTGCTGCGCACGTACCTCAAAGGTGCTCAAAGGTGCCGC 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 TTCGTGCTGCGCGTGGCGGGTCGTCTGCCTCAGGGTCGCCAGGGTCGGCAGC 600

Qy   601 TTCCATTGCAAGCGCATCGAGCTGATCACCATGCAGAGACGCTGGGTGCTGCGGGTGGACTG 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGCCCACTGCCAGCGCCATGGACACCGTCACCATGAAGCGCCTTCGCGCCCGCCACCTG 660

Qy   661 CACCCCAGTGTGCGGGAACGCTGTCTCGGAGGAGCAGAAGCCGGAGCGGACAACGAGGCACC 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 CACCCCAGTGTGGCGGCCCTGCCTCATCCAGCAGAAAGGCGGAGCGGGCACC 720

Qy   721 AAGAAGATCAGCACCTTCATAGGGACTTCCTTGTTGGTCTTCGGCCGCCCTATGTGATCACC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 AGGAAGATTGGCATTGCGACCTTCCTCATCCTGCTTTTGTTCGGCCGCCCTATGTGATCACC 780

Qy   781 AGGCTAGTTGGGAGGCTCTTCTCCACGGTGCCCATCGGCGCTCCCCACTGGGGGGTGTCCTGTGTCCAAG 840
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 AGGCTGGGCGGAGCTCGTGCCCTTCGTCTCACCGTGAACGCCCCAGTGGGGCATCCTCAGCAAG 840

Qy   841 TGCTTGGCGTACAGCAAGAGGCCGCATCCGACCCCTTGTGTTGTACTCCTTACTGCGACACCAG 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 TGCCTGACCTACAGCAAGAGGGGTGGGGCCGGTTCACGTACTCTTCTGCTGCCGTTCGGCCG 900

Qy   901 TTCCGCCAAAAAGCTGCAAGGAGAGATTCTGACACAGACGCTTCATCCACTCC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TACCGCAAAAGCTCGAAGGAGATTGTGCACAGACGACCCGGCCGCAGCA 960

Qy   961 TCTGGGCC 967
         ||  ||
Db   961 TCCACCC 967
```

Search completed: July 5, 2005, 16:12:07
Job time : 737 secs

OM nucleic - nucleic search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

Run on: July 5, 2005, 12:50:16 ; Search time 210 Seconds
               (without alignments)
               7900.879 Million cell updates/sec

Title: US-10-735-991-5
Perfect score: 1014
Sequence: 1 atgaactcgtgggacgcggg........ttctgccggtgtctgagtga 1014
Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0
Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents NA:*
               1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
               2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
               3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
               4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
               5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
               6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1009.2	99.5	2581	2	US-09-013-634-1	Sequence 1, Appli
2	1006.2	99.2	1011	2	US-09-013-634-3	Sequence 3, Appli
3	79	7.8	1047	3	US-08-540-650B-6	Sequence 6, Appli
4	79	7.8	1050	4	US-09-826-509-502	Sequence 502, App
5	79	7.8	1053	4	US-09-016-434-1423	Sequence 1423, Ap
6	79	7.8	1110	3	US-08-513-974B-31	Sequence 31, Appl
7	79	7.8	1110	4	US-08-776-971-26	Sequence 26, Appl
8	79	7.8	1110	4	US-09-461-436B-31	Sequence 31, Appl
9	79	7.8	1110	4	US-09-576-290-26	Sequence 26, Appl
10	79	7.8	1331	4	US-08-513-974B-322	Sequence 322, App
11	79	7.8	1331	3	US-08-776-971-103	Sequence 103, App
12	79	7.8	1331	4	US-09-576-290-103	Sequence 103, App
13	79	7.8	1882	3	US-08-540-650B-11	Sequence 11, Appl
14	79	7.8	3083	3	US-08-693-308-1	Sequence 1, Appl
15	73.8	7.3	1335	2	US-08-985-090-3	Sequence 3, Appli
16	73.8	7.3	1335	2	US-09-165-543-3	Sequence 3, Appli
17	73.8	7.3	1335	3	US-09-167-354-6	Sequence 6, Appl
18	73.8	7.3	1335	3	US-09-642-855-6	Sequence 6, Appl
19	73.8	7.3	1335	3	US-09-642-514-6	Sequence 6, Appl
20	73.8	7.3	2050	4	US-09-891-053-21	Sequence 21, Appl
21	73.8	7.3	2050	4	US-09-642-514-6	Sequence 6, Appl
22	73.8	7.3	2665	4	US-09-949-016-5059	Sequence 5059, Ap
23	73.8	7.3	2689	3	US-08-985-090-1	Sequence 1, Appl
24	73.8	7.3	2689	3	US-09-165-543-1	Sequence 1, Appl
25	73.8	7.3	2699	3	US-09-167-354-5	Sequence 5, Appli
26	73.8	7.3	2699	3	US-09-642-855-5	Sequence 5, Appli
27	73.6	7.3	1161	1	US-08-086-439C-2	Sequence 2, Appli
28	73.6	7.3	1161	3	US-08-434-877-2	Sequence 2, Appli
29	73.6	7.3	1367	3	US-08-475-742-3	Sequence 3, Appli
30	73.6	7.3	1367	4	US-08-261-293-3	Sequence 3, Appli
31	73.6	7.3	1504	4	US-09-016-434-1276	Sequence 1276, Ap
32	72	7.1	1370	4	US-08-056-051-1	Sequence 1, Appli
33	72	7.1	1370	1	US-07-928-611-17	Sequence 17, Appl
34	72	7.1	1370	1	US-08-487-811A-17	Sequence 17, Appl
35	72	7.1	1370	3	US-09-060-694-17	Sequence 17, Appl
36	72	7.1	1370	3	US-09-378-074-17	Sequence 17, Appl
37	72	7.1	1370	5	PCT-US93-07370-17	Sequence 17, Appl
38	72	7.1	1466	1	US-08-056-051-3	Sequence 3, Appl
39	72	7.1	1466	1	US-07-928-611-19	Sequence 19, Appl
40	72	7.1	1466	1	US-08-487-811A-19	Sequence 19, Appl
41	72	7.1	1466	3	US-09-060-694-19	Sequence 19, Appl
42	72	7.1	1466	3	US-09-378-074-19	Sequence 19, Appl
43	72	7.1	1466	5	PCT-US93-07370-19	Sequence 19, Appl
44	72	7.1	1610	1	US-08-056-051-5	Sequence 5, Appli
45	72	7.1	1610	1	US-07-928-611-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-013-634-1
Sequence 1, Application US/09013634
; Patent No. 5945307
; GENERAL INFORMATION:
; APPLICANT: M. Alexandra Glucksmann and Keith Robison
; TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,634
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Jean M. Silveri
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: MNI-036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 184..1194
US-09-013-634-1

Query Match   99.5%; Score 1009.2; DB 2; Length 2581;
Best Local Similarity   99.7%;
Matches 1011; Conservative   0; Mismatches   3; Indels   0; Gaps   0;
Pred. No. 3e-205;

```
RESULT 2
US-09-013-634-3
; Sequence 3, Application US/0901634
; Patent No. 5945307
```

GENERAL INFORMATION:
~   APPLICANT: M. Alexandra Glucksmann and Keith Robison
~   TITLE OF INVENTION: LIGAND RECEPTORS AND USES THEREFOR
~   NUMBER OF SEQUENCES: 14
~   CORRESPONDENCE ADDRESS:
~   ADDRESSEE: LAHIVE & COCKFIELD, LLP
~   STREET: 28 State Street
~   CITY: Boston
~   STATE: Massachusetts
~   COUNTRY: USA
~   ZIP: 02109
~   COMPUTER READABLE FORM:
~   MEDIUM TYPE: Floppy disk
~   COMPUTER: IBM PC compatible
~   OPERATING SYSTEM: PC-DOS/MS-DOS
~   SOFTWARE: Patentin Release #1.0, Version #1.25
~   CURRENT APPLICATION DATA:
~   APPLICATION NUMBER: US/09/013,634
~   FILING DATE:
~   CLASSIFICATION:
~   PRIOR APPLICATION DATA:
~   APPLICATION NUMBER:
~   FILING DATE:
~   ATTORNEY/AGENT INFORMATION:
~   NAME: Attorney, Jean M. Silveri
~   REGISTRATION NUMBER: 39,030
~   REFERENCE/DOCKET NUMBER: MNI-036
~   TELECOMMUNICATION INFORMATION:
~   TELEPHONE: (617)227-7400
~   TELEFAX: (617)742-4214
~   INFORMATION FOR SEQ ID NO: 3:
~   SEQUENCE CHARACTERISTICS:
~   LENGTH: 1011 base pairs
~   TYPE: nucleic acid
~   STRANDEDNESS: single
~   TOPOLOGY: linear
~   MOLECULE TYPE: cDNA
~   FEATURE:
~   NAME/KEY: CDS
~   LOCATION: 1..1011
~US-09-013-634-3

```
Query Match            99.2%;  Score 1006.2; DB 2; Length 1011;
Best Local Similarity  99.7%;  Pred. No. 1.1e-204;
Matches 1008; Conservative  0; Mismatches  3; Indels  0; Gaps  0;
```

RESULT 3
US-08-540-650B-6
; Sequence 6, Application US/08540650B
; Patent No. 6399325
; GENERAL INFORMATION:
;   APPLICANT: HINUMA, Shuji
;   APPLICANT: FUJII, Ryo
;   APPLICANT: FUKUSUMI, Shoji
;   APPLICANT: OHTAKI, Tetsuya
;   APPLICANT: HOSOYA, Masaki
;   APPLICANT: OHGI, Kazuhiro
;   APPLICANT: ONDA, Haruo
;   TITLE OF INVENTION: GALANIN RECEPTOR PROTEIN, PRODUCTION AND USE THEREOF
;   NUMBER OF SEQUENCES: 17
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
;   STREET: 130 WATER STREET
;   CITY: BOSTON
;   STATE: MA
;   COUNTRY: US
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ Version 1.5
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/540,650B
;   FILING DATE: 11-OCT-1995
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 7-134412
;   FILING DATE: 31-MAY-1995
;   APPLICATION NUMBER: 6-326610
;   FILING DATE: 28-DEC-1994
;   APPLICATION NUMBER: 6-247599
;   FILING DATE: 13-OCT-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: RESNICK, DAVID S
;   REGISTRATION NUMBER: 34,235
;   REFERENCE/DOCKET NUMBER: 45901
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-523-3400
;   TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1047
;   TYPE: Nucleic acid
;   STRANDEDNESS: Double
;   TOPOLOGY: Linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   IDENTIFICATION METHOD: S
US-08-540-650B-6

Query Match   7.8%; Score 79; DB 3; Length 1047;
Best Local Similarity 49.3%; Pred. No. 9.5e-08;
Matches 264; Conservative   0; Mismatches 265; Indels   6; Gaps   2;

```
Qy   59 TGCTGTCCAACGCGCGTGGGTGCTGGTGCTTGCCTTCTGCGCTGCCACAGC---GCGGACAATCCGGCGCC 115
         ||||||||||||||||||| ||||| |||| | ||||||| |||   ||||||   |||||| ||||| |||
Db  143 TGCTGGGCAACAGCCTAGTGATCACGTGTTGGGGCGGGCGCGCAGAAGCCGGGCAAGCCGCGGA         202

Qy  116 AGGGCGCCGGCGCGCTCTTCACCCTGAACCTCAGGTGCGGGGAAACCTGCTGTGTGCACCGTGGGTCA     175
         | ||||||||||||||||| || |||||||||| ||||||||||||||||||| |||||||||||||||
Db  203 GCACCACCAACCTGTTCATCCTCAACCTGAGCATCGCGACCTGGGCTACCTGCTCTTCT             262

Qy  176 ACATGCGCGTCACGTGGCCGGCCGGTCGTGGGGCGCAGGCGGCCAGCGACCGGCGCTGT             235
         ||| |||||||| || ||||||||| |||||||||| ||||||||||  |||||||| ||
Db  263 GCATCCCCTTCCAGGCCGACCGGTGTACGCGCGTGCCCACCTGGGGTGCTCGGGCGCCTTCATCT       322

Qy  236 GCCGGCCTGGGCTGTGCCCTTCCTCGACACCTTCCTGGGTGCCAACTCCATGCTCAGCATGGCCG      295
         |||||||||||| |  ||||  |||||| |||||||||| || ||||| ||||| |||| ||||||
Db  323 GCAAGTTCATCCACTACTTCTTCACCGTGTCCATGCTGGGTGAGCATCTTCACCCTGCCG         382
```

RESULT 4
US-09-826-509-502
; Sequence 502, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
;   APPLICANT: Lehmann-Bruinsma, Karin
;   APPLICANT: Liaw, Chen W.
;   APPLICANT: Lin, I-Lin

```
Qy  296 CGCTCAGCATCGACCCGCTGGGGTGGCCGTGGGTGCTTCCCGCTACGGGGCCGAAGATGC 355
         |||||||||||||| |||| ||||||||||||| || ||||||||||||| ||||||||||
Db  383 CGATGTCCGTGGACCCGCTACGGTGCACTCGCGCGCGCTCCTCTCCCTCAGGG       442

Qy  356 GCCTCCGGACGCGGGCGGGCTCATGGTGGCGCCTACACGTGGCGTCACGCGGCTCACCTTCCCAG 415
         ||||||||||||| ||||||||| |||||| ||||| |||||||| | |||| ||||||||||||
Db  443 TGTCCCGCAACGGCGCTGCTGGGCGTGGGCGTGGCGTCCATCTGGGGCGTCCATTGGCCAT---GG 499

Qy  416 CCGCCGCGGGCTGGCCCCTGTCCTGGCCGTTCGGGCTTCCACCAGCGTGACGCTGCTGT 475
         |||||||||||||||| ||| ||||||||||| | ||| ||||||||||||||||| |||
Db  500 CCTCGCCGTGGCCTACCACCAGGGGCCTCTTCCACCGCGCGCGCCAACAGACCTTCT       559

Qy  476 GCAGCCCGGCGGCCAGACGAGGCCCTGGCGTTCGGCCGTTCACTGGCGCCTTCCAGGCTC 535
         |||| ||||||||| |||||||||| |||||||| |||||||| || || ||| ||| |||
Db  560 GCTGGGAGCAGTGGCCGGACCCGCCACAAGAAGGCCTACGTGGTGTGGGCAACCTTCGTCT 619

Qy  536 TCAGCTTCCTGCTCTCCTTGTCTGCTGCTCGCGACGTAGCTCAAGGTGCTCAA 590
         ||||| ||| |||| ||||| |||| |||||||| ||||||||| || ||| ||
Db  620 TCGGCTTACCTGCTGCGGCCTCCTCACTTCTCGCTATGCCAAGGTGCTCCTTAA 674
```

; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; TITLE OF INVENTION: Protein-Coupled Receptors
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
;
; SEQ ID NO 502
; LENGTH: 1050
; TYPE: DNA
; ORGANISM: Homo sapiens
;
US-09-826-509-502

Query Match 7.8%; Score 79; DB 4; Length 1050;
Best Local Similarity 49.3%; Pred. No. 9.5e-08;
Matches 264; Conservative 0; Mismatches 265; Indels 6; Gaps 2;

```
Qy   59 TGCTGTCCAACGCGCGCTGGTGCTGCTGCTGCTGCTGCTGCTGCACAGC---GCGGACATCCGCCGCC 115
        ||||       |||     ||||  ||    |     ||||||||        ||||||  || |||
Db  143 TGCTGGGCAACAGCCTAGTGAATCACCGTGCTGGGCGGGCGCAGCGAAGCCGCGGGA 202

Qy  116 AGGCGCCCGGCGGGCTCTTCACCCTGAACCTCACGTGGGGGAACCTGCTGTGTGACCGTGGGTCA 175
        |||     |||||  |||||    ||||||||    ||  ||||   ||||||||||||  ||||
Db  203 GCACCACCAACCTGTTCATCCTCAACCTGAGCATCGCCGACCTGGCCTACCTGCTCTTCT 262

Qy  176 ACATGCCGCTCACGCTGGCCGCGGGCTCGTGTGGCCGCAGCGGCCGGACCGGCCTGT 235
        |||||| ||||||   ||||  ||  ||   | |||| ||   |||    ||||||
Db  263 GCATCCCTTCCAGGCCACCGGTGTACGGCGTGCCACCTGGGGCGTCTGGGGCCTTCATCT 322

Qy  236 GCCGGCCTGGCTGCCTTCCTCGACACCTTCCTGCTCAGCATGGCCG 295
        |||  | |||||||||  || |||||  ||||  |  | |||   |
Db  323 GCAAGTTCATCCACTACTTCTTCACGGCGTCAGCATGGCCG 382

Qy  296 CGCTCAGCATCGACGCGGTGGTGCTTCGTCTTCCCGCTGTAGCGGCCGCAAGATGC 355
        || |||||  ||||  | || ||| | ||   ||||  |  ||||||  |||||
Db  383 CGATGTCCGTGGACCCGTACCTGGCCATCGTGCACCTGCCTCCCTCAGGG 442

Qy  356 GCCTCCGCGACGGGCGGGGGCTCATGGGTGGCCTACACGTGGCCGCAGCGGGC 415
        ||||||  ||||||||||| ||   |  | || ||| ||||  |||||  |||
Db  443 TGTCCCGCACGCGCTGGGCGGGCGCTGCATCTGGGGCGGCGTGCCATCGTCCATTGCCAT 499

Qy  416 CCGCCGCCTCGCCCTGTCCTTCGCCGGCCGCTGGGGCGCTTCGCTC 475
        |||||  |||||    ||| ||||||||| |  ||||| ||  |
Db  500 CCTCGCCCGTGGCCTTACCACCAGGGCCTTCCACGGGG 535

Qy  476 GCAGCGCGGCGGCCAGACGACGACGCCTGCGCTTCGCCGTCTTCACTGGCGCGCTC 535
        ||||||  ||||||  | |||||   |||  |||| |||  ||||||  |||||
Db  536 TCAGCTTCTCGCTCTCCTTCGTGTTCGTTCGTGCACGTACCTCAAGGTGCTCAA 590

Qy  536 TCAGCTTCTCGCTCTCCTTCGTGTTCGTTCGTGCACGTACCTCAAGGTGCTCAA 590
        |||||  ||  | ||||||||||  ||| ||  || |||   ||||| ||| |
Db  560 GCTGGGAGCAGTGGGCCCGACCCTACGTGGTGTGCACCTTCGTCT 619

Qy  560 GCTGGGAGCAGTGGGCCCGACCCTACGTGGTGTGCACCTTCGTCT 619
        |||||      |  | |||  | | |  |||| |  |||||
Db  620 TCGGCTTACCTGCTGCGCGCCTCCTGCTCATCTGCTTCTGCTATGCCAAGGTCCTTAA 674
```

RESULT 5
US-09-016-434-1423
; Sequence 1423, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1423:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1053 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9555047
US-09-016-434-1423

Query Match 7.8%; Score 79; DB 4; Length 1053;
Best Local Similarity 49.3%; Pred. No. 9.5e-08;
Matches 264; Conservative 0; Mismatches 265; Indels 6; Gaps 2;

```
Qy   59 TGCTGTCCAACGCGCGCTGGTGCTGCTGCTGCTGCTGCTGCTGCACAGC---GCGGACATCCGCCGCC 115
        ||||       |||     ||||  ||    |     ||||||||        ||||||  || |||
Db  143 TGCTGGGCAACAGCCTAGTGAATCACCGTGCTGGGCGGGCGCAGCGAAGCCGCGGGA 202

Qy  116 AGGCGCCCGGCGGGCTCTTCACCCTGAACCTCACGTGGGGGAACCTGCTGTGTGACCGTGGGTCA 175
        |||     |||||  |||||    ||||||||    ||  ||||   ||||||||||||  ||||
Db  203 GCACCACCAACCTGTTCATCCTCAACCTGAGCATCGCCGACCTGGCCTACCTGCTCTTCT 262

Qy  176 ACATGCCGCTCACGCTGGCCGCGGGCTCGTGTGGCCGCAGCGGCCGGACCGGCCTGT 235
        |||||| ||||||   ||||  ||  ||   | |||| ||   |||    ||||||
Db  263 GCATCCCTTCCAGGCCACCGGTGTACGGCGTGCCACCTGGGGCGTCTGGGGCCTTCATCT 322

Qy  236 GCCGGCCTGGCTGCCTTCCTCGACACCTTCCTGCTCAGCATGGCCG 295
        |||  | |||||||||  || |||||  ||||  |  | |||   |
Db  323 GCAAGTTCATCCACTACTTCTTCACGGCGTCAGCATGGCCG 382

Qy  296 CGCTCAGCATCGACGCGGTGGTGCTTCGTCTTCCCGCTGTAGCGGCCGCAAGATGC 355
        || |||||  ||||  | || ||| | ||   ||||  |  ||||||  |||||
Db  383 CGATGTCCGTGGACCCGTACCTGGCCATCGTGCACCTGCCTCCCTCAGGG 442

Qy  356 GCCTCCGCGACGGGCGGGGGCTCATGGGTGGCCTACACGTGGCCGCAGCGGGC 415
        ||||||  ||||||||||| ||   |  | || ||| ||||  |||||  |||
Db  443 TGTCCCGCACGCGCTGGGCGGGCGCTGCATCTGGGGCGGCGTGCCATCGTCCATTGCCAT 499

Qy  416 CCGCCGCCTCGCCCTGTCCTTCGCCGGCCGCTGGGGCGCTTCGCTC 475
        |||||  |||||    ||| ||||||||| |  ||||| ||  |
Db  500 CCTCGCCCGTGGCCTTACCACCAGGGCCTTCCACGGGG 535

Qy  476 GCAGCGCGGCGGCCAGACGACGACGCCTGCGCTTCGCCGTCTTCACTGGCGCGCTC 535
        ||||||  ||||||  | |||||   |||  |||| |||  ||||||  |||||
Db  536 TCAGCTTCTCGCTCTCCTTCGTGTTCGTTCGTGCACGTACCTCAAGGTGCTCAA 590

Qy  536 TCAGCTTCTCGCTCTCCTTCGTGTTCGTTCGTGCACGTACCTCAAGGTGCTCAA 590
        |||||  ||  | ||||||||||  ||| ||  || |||   ||||| ||| |
Db  560 GCTGGGAGCAGTGGGCCCGACCCTACGTGGTGTGCACCTTCGTCT 619

Qy  560 GCTGGGAGCAGTGGGCCCGACCCTACGTGGTGTGCACCTTCGTCT 619
        |||||      |  | |||  | | |  |||| |  |||||
Db  620 TCGGCTTACCTGCTGCGCGCCTCCTGCTCATCTGCTTCTGCTATGCCAAGGTCCTTAA 674
```

RESULT 6
US-08-513-974B-31
; Sequence 31, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
;   APPLICANT: Hinuma, Shuji
;   APPLICANT: Hosoya, Masaki
;   APPLICANT: Fujii, Ryo
;   APPLICANT: Ohtaki, Tetsuya
;   APPLICANT: Fukusumi, Shoji
;   APPLICANT: Ohgi, Kazuhiro
;   TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
;   PRODUCTION, AND USE THEREOF
;   NUMBER OF SEQUENCES: 380
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;   STREET: 130 Water Street
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/513,974B
;   FILING DATE: 14-SEP-1995
;   CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/JP95/01599
;   FILING DATE: 10-AUG-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 7-093989
;   FILING DATE: 19-AUG-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 7-057186
;   FILING DATE: 16-MAR-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 7-007177
;   FILING DATE: 20-JAN-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 6-326611
;   FILING DATE: 28-DEC-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 6-270017
;   FILING DATE: 02-NOV-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 6-236357
;   FILING DATE: 30-SEP-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 6-236356
;   FILING DATE: 30-SEP-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 6-189274
;   FILING DATE: 11-AUG-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 6-189273
;   FILING DATE: 11-AUG-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 6-1945
;   FILING DATE: 11-AUG-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: JP 6-189272
;   FILING DATE: 11-AUG-1994
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Resnick, David S.
;   REGISTRATION NUMBER: 34,235
;   REFERENCE/DOCKET NUMBER: 45753
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-523-3400
;   TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 31:
;   SEQUENCE CHARACTERISTICS:

;   LENGTH: 1110 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-513-974B-31

Query Match
Best Local Similarity   7.8%;   Score 79;   DB 3;   Length 1110;
Matches 212; Conservative 51.1%; Pred. No. 9.6e-08; Mismatches 200; Indels 3; Gaps 1;

```
Qy   18 GGGGCCTGCGCGGCGGGGGCTRACTGTGGGCACGACGATGGGGGTTCCTGCGTGCTGTGTTCCAACGCGCTGGT   77
          ||||||   |     |       |    ||    | ||   |      ||   |
Db  180 GGGGCGTGATCGTGCTGCTCTCTACAGCGTCGGTGGTGGTGGTCGTGTGGGGCTGGTGGGGCAACTGCCT     239

Qy   78 GCTGCTCTGCCTGCTGCGCACAGCGCGGACAATCCGCGCCGGCCGCCAGGCGCGCCGGCGCGCTCTTCACCCT  137
         |||| |||| || ||||| |||  |    |    ||   |    | |
Db  240 GCTGGTGCTGGTGATCGGCGGCGGGGGTGCGCCGGCCTGCACAACGTGACGAACTTCCTCATCGG          299

Qy  138 GAAACCTCACGTGCGGGGAACCTGCTGTGCAGCCGTGGTCTAAACATGCCCGCTCAGCGTGCCGG           197
         |||||| | |  |      |   ||    | |    |   ||     |||||||    ||||
Db  300 CAAACCTGCCCTTGTCCGACGTGCCTCATGTGCACCGCGCCGCGCTCAGCGCCTGCCCTA           359

Qy  198 CGTCGTGGGCGCCAGGCGCGCCAGCCGCGCGG---GGGGACGCGCCTGTGCCGCCTGGCGCGCCTTCCT      254
         ||||| | |   ||    |  | |||  |      |||  |     ||    |   ||    |
Db  360 TGCCTTCGAGCCACGCGGGCGTGGGTGTTCGGCGGCGGCGGCCTGTGCCACCTGGTCTTCTTCCT          419

Qy  255 CGACACCTTCCTGGCTGCCAACTCAGCAATGGCCGCGGCTCAGCAATCGACCGCTG                   314
         |    ||| | ||    | |  | || |   | |    || | |     ||||||
Db  420 GCAGCCGGTCACCGTCTATGTGTCGGTGTTCACGCTCACCATCGGCAGCATGGGACGCGTA               479

Qy  315 GGTGGCGCGTGGTCTTCCCGCTGAGCTACCGGGCCAAGAATGGGCCTCCGCGACGGCGGCGCT              374
         || | | || || |  |||||  ||    |  |   | ||  |   | || | ||
Db  480 CGTCGTGCTGGTGCACCCGTTGAGGGCGGGGGGCCATCGCGTGCGCCCTCAGGCGGCCTACCGGTGT        539

Qy  375 CATGGTGGGCCTTACACGTGGGCTGCACGCGGCTCACCTTCCCAGCCGCGCGCGCGGCGTTGGCC            429
         | || ||||  | |  | |     |    | ||     |       |   |  |||  |
Db  540 GCTGGCCATCTGGGCGCGGTGCCGCGGCTGTTCGGGCCGGTGTTGCCCCGGCCGCGTGCACACC           594
```

RESULT 7
US-08-776-971-26
; Sequence 26, Application US/08769971B
; Patent No. 6228984
; GENERAL INFORMATION:
;   APPLICANT: Hinuma, Shuji
;   APPLICANT: Habata, Yugo
;   APPLICANT: Kawamata, Yuji
;   APPLICANT: Hosoya, Masaki
;   APPLICANT: Fujii, Ryo
;   APPLICANT: Fukusumi, Shoji
;   APPLICANT: Kitada, Chieko
;   TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
;   NUMBER OF SEQUENCES: 140
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;   STREET: 130 Water Street
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02109
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/776,971B
;   FILING DATE: 06-Feb-1997
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/JP96/03821
;   FILING DATE: 28-DEC-1996

Left column

```
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 28-DEC-1995
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; APPLICATION NUMBER: JP 8/211805
; FILING DATE: 12-AUG-1996
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME: Conlin, David G.
;   REGISTRATION NUMBER: 27,026
;   REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-523-3400
;   TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 26:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1110 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-08-776-971-26
```

```
Query Match             51.1%;  Score 79;  DB 3;  Length 1110;
Best Local Similarity
Matches 212;  Conservative  7.8%;  Pred. No. 9.6e-08;  Mismatches 200;  Indels  3;  Gaps  1;

Qy    18  GGGGCCTGGGCGGGCGGGGCTACTGGTGGGCACGAGATGGGGCGTCTCGCTGCTGTTCCAACGCGCTGGT   77
          |||||  |                                   |  |||             |||||||
Db   180  GGGGGCTGATCGTGCGTCCTCTACAGCCGTTCGGGCTGGTGGGCAACTGCCT  239

Qy    78  GCTGCTCTGCCTGCTGCTGACACAGCGCGGACATCCGCGCGCCGAGCGCGCCGGGCTCTTCACCCT  137
          |||| |    |  |  |   ||||     ||  |||||||  ||||| ||   ||||  |||||
Db   240  GCTGGTGCTGGTGATCGGCGCGGGGTTGCCCCCGGGCTGCAACAACCGTGACGAACTTCCTCATCGG  299

Qy   138  GAACCTCACGTGCGGGGAACCTGCTGTGTCACGTGTGTCAACATGCCGCCTCACGCTGGCCGG  197
          |  |||      ||  ||  |||| || |  ||||||||  ||  |      |||||||||||
Db   300  CAACCTGGCCTTGTCCGACGTGCCTCATGTGCACCCGTGCCGCCTCACGCGCTGGCCTA  359

Qy   198  CGTCGTGCGCGCAGCGGCAGCGGCGCCG---GGGCACCCGCCTCGTGCCGCCTGGCTTCCT  254
          |||  |                         |||  |||     |    || ||||| ||||
Db   360  TGCCTTCGAGCCACCGGCGGCTGGGTGTTCGGCGGCGGCCTGTGCCACTCGGTTCTTCTTCCT  419

Qy   255  CGACACCTTCCTGGGTGCCAACTCCATGCTCGTCAGCGATGGCATCATCGACCGGCTG  314
          ||  |  |       |||||  |    |   |       ||||    ||||||| |||
Db   420  GCAGCCGGTCACCGGTCATAGTGTCGGGTGTTCAGCCTCACCAGTGGACCGGCTA  479

Qy   315  GGTGGCCGTGGTGTCTTCCCGGCTGGAGCTACCGGGGCCAAGATCGGCCTTCCGCGACGCGCGGCGCT  374
          |||||| |  ||| |      ||  |    |  |     |||||||  ||  |  |||  |   |||||
Db   480  CGTCGTGCTGCTGTGCACCGGCCATCTCGCTCTCGCTCTCAGCGCGGCCTTACGGCGGGCGCT  539

Qy   375  CATGGTGGCCTACACGTGGCTGCACGCGCTCACCTTCCAGCCGCCGCCGGCTCGCC  429
          |||| |||||       ||  |   |  |  |  |     |  |      |  |||
Db   540  GCTGGGCATCTGGGGCTGGCCTGGTCCGCGGCGGCGCCATCTGGCTTCAGGCGCTCCGGTGCACC  594
```

Right column

```
; STATE: MA
; COUNTRY: USA
; ZIP: 02209
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/461,436B
;   FILING DATE: 14-Dec-1999
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/513,974
;   FILING DATE: 14-SEP-1995
;   APPLICATION NUMBER: PCT/JP95/01599
;   FILING DATE: 10-AUG-1995
;   APPLICATION NUMBER: 7-033989
;   FILING DATE: 11-AUG-1995
;   APPLICATION NUMBER: 6-189274
;   FILING DATE: 19-APR-1995
;   APPLICATION NUMBER: 6-189273
;   FILING DATE: 11-AUG-1994
;   APPLICATION NUMBER: 6-189272
;   FILING DATE: 11-AUG-1994
;   APPLICATION NUMBER: 7-057186
;   FILING DATE: 16-MAR-1995
;   APPLICATION NUMBER: 7-007177
;   FILING DATE: 20-JAN-1995
;   APPLICATION NUMBER: 6-326611
;   FILING DATE: 28-DEC-1994
;   APPLICATION NUMBER: 6-270017
;   FILING DATE: 02-NOV-1994
;   APPLICATION NUMBER: 6-236357
;   FILING DATE: 30-SEP-1994
;   APPLICATION NUMBER: 6-236356
;   FILING DATE: 30-SEP-1994
; ATTORNEY/AGENT INFORMATION:
;   NAME: CONLIN, DAVID G.
;   REGISTRATION NUMBER: <Unknown>
;   REFERENCE/DOCKET NUMBER: 45753 DIV2
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-439-4444
;   TELEFAX: 617-439-4170
; INFORMATION FOR SEQ ID NO: 31:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1110 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-461-436B-31
```

```
Query Match             51.1%;  Score 79;  DB 4;  Length 1110;
Best Local Similarity
Matches 212;  Conservative  7.8%;  Pred. No. 9.6e-08;  Mismatches 200;  Indels  3;  Gaps  1;

Qy    18  GGGGCCTGGGCGGGCGGGGCTACTGGTGGGCACGAGATGGGGCGTCTCGCTGCTGTTCCAACGCGCTGGT   77
          |||||  |                                   |  |||             |||||||
Db   180  GGGGGCTGATCGTGCGTCCTCTACAGCCGTTCGGGCTGGTGGGCAACTGCCT  239

Qy    78  GCTGCTCTGCCTGCTGCTGACACAGCGCGGACATCCGCGCGCCGAGCGCGCCGGGCTCTTCACCCT  137
          |||| |    |  |  |   ||||     ||  |||||||  ||||| ||   ||||  |||||
Db   240  GCTGGTGCTGGTGATCGGCGCGGGGTTGCCCCCGGGCTGCAACAACCGTGACGAACTTCCTCATCGG  299

Qy   138  GAACCTCACGTGCGGGGAACCTGCTGTGTCACGTGTGTCAACATGCCGCCTCACGCTGGCCGG  197
          |  |||      ||  ||  |||| || |  ||||||||  ||  |      |||||||||||
Db   300  CAACCTGGCCTTGTCCGACGTGCCTCATGTGCACCCGTGCCGCCTCACGCGCTGGCCTA  359

Qy   198  CGTCGTGCGCGCAGCGGCAGCGGCGCCG---GGGCACCCGCCTCGTGCCGCCTGGCTTCCT  254
          |||  |                         |||  |||     |    || ||||| ||||
Db   360  TGCCTTCGAGCCACCGGCGGCTGGGTGTTCGGCGGCGGCCTGTGCCACTCGGTTCTTCTTCCT  419
```

RESULT 9
US-09-576-290-26
; Sequence 26, Application US/09576290
; Patent No. 6794491
; GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Habata, Yugo
APPLICANT: Kawamata, Yuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Fukusumi, Shoji
APPLICANT: Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/576,290
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/776,971
FILING DATE:
PRIOR APPLICATION DATA:
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/211805
PRIOR APPLICATION DATA:
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/59419
PRIOR APPLICATION DATA:
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 7/343371
PRIOR APPLICATION DATA:
FILING DATE: 18-SEP-1996
APPLICATION NUMBER: JP 8/246573
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1110 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA

Qy 255 CGACACCTTCCTGGCTGCCAACTCCATGCTCAGCATGGCCCCGCGCTCAGCACTCGGCTG 314
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 420 GCAGCCGGTCACCGTCTATGTCCGGTGTTCACGGCTCACCACCATGCCAGTGGACCGGCTA 479

Qy 315 GGTGGCCGGTGGTCTTCCCGCTGAGCTACCGGCCCAAGATCGCGCTTCCGGACAGCGGCGCGCT 374
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 480 CGTCGTGCTGCTGCACCGGCTGAGGCGGGGCGCATCTGCTGCTCCAGCGCGCTACCGCTGT 539

Qy 375 CATGGTGGCCTACACGTGGCGTGCACCGGCTTCAACCTTCCCAGCCGCTCGGCC 429
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 540 GCTGCGGCCATCTGGGCGCGTCCGCGCGGCTGCTTGGCGGCCGCCGGTGCACACC 594

US-09-576-290-26
Query Match 212; Best Local Similarity 51.1%; Score 79; DB 4; Length 1110;
Matches 212; Conservative 0; Mismatches 200; Pred. No. 9.6e-08; Indels 3; Gaps 1;

Qy   18 GGGCCTGGCGGGGGGCCTACTGGTTGGGCACGAATGGGCGTCGGTGCTGCCTTGGT 77
         ||| ||| ||||| ||| ||||||| |||| ||||||||||| |||||||||||
Db 180 GGGGCTTGAATCGTTGCTGCTCTCACAGCCGTCGTCGGTTGGTGGGCAAACTGCCT 239

Qy   78 GCTGCTCTTGCCCTGCTGCACAGCGCGGACATCGCGCGCCGCAGGGCGGCCGCCCT 137
         || ||| ||||||||||||||||||||||||| ||||||||| |||||||||||||
Db 240 GCTGGTGCTGGTTGATCGCGCGCGGGTGCCCGGCGTGCACGAACTTCCTCATCGG 299

Qy 138 GGAACCTCACGTGCGGGAAACCTGCTGTGCACCGTGGTCAAACATGCCCGTCACGGCTCAGCGCGG 197
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db 300 CAACCTGGCCTTGTCCGACGTGCTCAGTGTGCACCGCGCTGCGGTGCCGCGCTCACCGCTGGGCTA 359

Qy 198 CGTCGTGCGCGAGCGGCCAGCGCGGCGGCGCGCG---GGGGACCGGCGCTGGCCGCCTTCCT 254
         |||||||||||||||||||||||||||| |||||| |||||| ||||||||||||||||
Db 360 TGCCTTCGAGCGCCACGCGGCGGTGTTCGGCGCGGGGCGCCGCCTTGCCACCTGGTCTTCTTCCT 419

Qy 255 CGACACCTTCCTGGCTGCCAACTCCATGCTCAGCATGGCCCCGCGCTGCAGCATCGACCGCTG 314
         ||| |||||||| |||||||||||||||||||||||||||||||| |||| |||||||||||
Db 420 GCAGCCGGTCACCGTCTATGTCGGTGTTCACGGCTCACCACCATGCCAGTGGACCGGCTA 479

Qy 315 GGTGGCCGGTGGTCTTCCCGCTGAGCTACCGGCCCAAGATCGCGCTTCCGGACAGCGGCGCGCT 374
         |||||||| || |||||||||||||||||| ||||||| |||||||||||||||| |||||
Db 480 CGTCGTGCTGCTGCACCGGCTGAGGCGGGGCGCATCTGCTGCTCCAGCGCGCTACCGCTGT 539

Qy 375 CATGGTGGCCTACACGTGGCGTGCACCGGCTTCAACCTTCCCAGCCGCTCGGCC 429
         ||| ||||||||||||||||||||||||||||||||||||||||||
Db 540 GCTGCGGCCATCTGGGCGCGTCCGCGCGGCTGCTTGGCGGCCGCCGGTGCACACC 594

RESULT 10
US-08-513-974B-322
; Sequence 322, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ohtaki, Tetsuya
APPLICANT: Fukusumi, Shoji
APPLICANT: Ohgi, Kazuhiro
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
TITLE OF INVENTION: PRODUCTION, AND USE THEREOF
NUMBER OF SEQUENCES: 380
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186

; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-007177
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-326611
; FILING DATE: 28-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236357
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189273
; FILING DATE: 11-AUG-1945
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189274
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189272
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 45753
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 322:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1331 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 118..1227

US-08-513-974B-322

Query Match   7.8%; Score 79; DB 3; Length 1331;
Best Local Similarity 51.1%; Pred. No. 9.9e-08;
Matches 212; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

```
Qy   18  GGGGCCTGGGCGGGGGCTACTGGTGGGGCGACGATGGGCGTCCGGCGCCGGTGCTGGG   77
          ||||   |||          |      ||      ||  |   ||||  ||
Db  297  GGGGGCTTGATCGGTCGTTCACGCGTTCTACAGCGCTCGTTGGTCGGGCT          356

Qy   78  GCTGCTCGTGCCTGCTGCACAGCGCGGACATCGCGCCCAGCGCCGCGGCGCTCTTCACCCT   137
Db  357  GCTCGGTGCTCGGTGATCGCGCGGGCGGCCCCGCGCCTGCACAAGCGTGACGAAACTTCCTCATCGG   416

Qy  138  GAACCTCACGTGCGCGGGAACCTGCTCGTGTGCACACGTGGTCAACATGCCGCTCAGCGCTGGCCGG   197
Db  417  CAACCTGGCGCCTTGTCCGACGTGCTCACNTGCCACCGCCTGCCTGCCCCTCACGCTGCCCTA   476

Qy  198  CGTCGGTGGGCGCGGCGCAGCCGGCG---GGGGACCGCGCCTGTGCCGGGCGCCGCCCTTCCT   254
Db  477  TGCCTTCGGAGGCCACGCGGGGTGTTGCGGGGCCTTGCGGTCCGCCACCTGGTGCTCTTTTCCT   536

Qy  255  CGACACCTTCCTGGCTGCCAACTCCATGCTTCAGCANTGCAATCGACCGGCTG   314
Db  537  GCAGCCGGTCACCGTCTATGTGTTCGCGTTCACCACCATCGACATGGGACACGGCTA   596

Qy  315  GGGTGGGCCGTGGTGCTTCTTCCCGCTGAGTCATACCGGGGCACGGATGGGCGTCTCGCTCGGTGTCCTTGCGGCTGGT   374
Db  537  GCAGCCGGTCACCGTCTATGTGTTCGCGTTCACCACCATCGACATGGGACACGGCTA   596

Qy  375  CATGGTGGCCTTACACGTGGTCGCACGCGCTCACCTTCCCAGCCGCGGGCGCTCGCC   429
Db  597  CGTCCGTTGCGTCGTGCACCCGGCGCATCTGCGCTTCGCTCTGCCCTTCAGCGCCTACGGTGT   656
```

Db 657 GCTGCGCCATCTGGGCGCCGGTGCTTCCGGCGCGTGCCGCCGGGCGTGGGCGCTGGGCGCTGGGCGCCGCCGTGCACACC   711

RESULT 11
US-08-776-971-103
Sequence 103, Application US/08776971B
; Patent No. 6228984
; GENERAL INFORMATION:
; APPLICANT:
;   Hinuma, Shuji
;   Habata, Yugo
;   Kawamata, Yuji
;   Hosoya, Masaki
;   Fujii, Ryo
;   Fukusumi, Shoji
;   Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
;   STREET: 130 Water Street
;   CITY: Boston
;   STATE: MA
;   COUNTRY: USA
;   ZIP: 02109
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/776,971B
;   FILING DATE: 06-Feb-1997
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/JP96/03821
;   FILING DATE: 28-DEC-1996
;   APPLICATION NUMBER: JP 7/343371
;   FILING DATE: 28-DEC-1995
;   APPLICATION NUMBER: JP 8/59419
;   FILING DATE: 15-MAR-1996
;   APPLICATION NUMBER: JP 8/211805
;   FILING DATE: 12-AUG-1996
;   APPLICATION NUMBER: JP 8/246573
;   FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
;   NAME: Conlin, David G.
;   REGISTRATION NUMBER: 27,026
;   REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-523-3400
;   TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1331 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; FEATURE:
;   NAME/KEY: Coding Sequence
;   LOCATION: 118..1227
; OTHER INFORMATION:
; SEQUENCE DESCRIPTION: SEQ ID NO: 103:

US-08-776-971-103

Query Match   7.8%; Score 79; DB 3; Length 1331;
Best Local Similarity 51.1%; Pred. No. 9.9e-08;
Matches 212; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

```
Qy   18  GGGGCCTGGGCGGGGGCTACTGGTGGGGCACGACGATGGGCGTTCGGCGTCCGTGTCCAAGCGGCTGGT   77
          ||||   |||          |      |      ||||  ||       ||  ||
Db  297  GGGGGCTTGATCGGTCGTTCGTCGTCTCTACAGCGCTCGTTGGTTGGGGCTCGGTGGGGGCAACTGCCT   356
```

Left column:

```
Qy   78  GCTGCTCTGCCTGCTCTGCACAGGCGGGGACATCCGCCGCCGCGGCGGCTCTTCACCCT  137
         |||||  ||||  |||   || ||   ||||   |||  |||  |||  ||| ||||
Db  357  GCTCGGTGCTCGGTGATCGGCGCGGGTGCGCCGCGGCTGACGGAACTTCCTCATCGG   416

Qy  138  GAACCTCACGTGCGGGGAACCTGCTGGTCACACATGCCGCTCACGCTGGCCGG       197
         ||||| ||  ||  |||  |||||||  |   ||| || |||||||||||||||
Db  417  CAACCTGGCCTTGTCCGACGTGCTCATGTGCACCGCCGCCTCAGCGCTGGCCTA      476

Qy  198  CGTCGTGGGGGCCAGCGGAGCCGGGG---GGCGACGCGCCTGTGCCGGCTGCCTTCCT  254
         |||||  ||| |||   ||  ||||   |||||| || |||||| ||  |||||||||
Db  477  TGCCTTCGAGCCACGGGGCGTGTTCGGCGGGGCGGGCCCTGTTGCCACCTGGTCTTCTTCCT  536

Qy  255  CGACACCTTCCTGGCTGCCAACTCCATGCTCAGCAATGGCCCGCGCTCAGCATCGACCGGCTG  314
         |||||||||||  ||||||||    |||   ||| ||| || || |||||||| || ||||||
Db  537  GCAGCCGGTCACCGTCTATGTGTTCACGGTTCACCTGGTCTTCTTCCT             596

Qy  315  GGTGGCCGTGGTCTTCCCGCTGAGCTACCGGGCCAAGATCGGCCTCCGAGCGCGGCGGCT  374
         ||||  ||| |||  ||| ||||||  ||  |||  |||  ||  | ||| || || |||
Db  597  CGTCCGTGCTGGTGCACCGGCCCATCTCGCTCGGGCCTCAGCGCCTACGGCTGT       656

Qy  375  CATGGTGGCCTACACGTGGGCTGCACGGCGCTCACCTTCCCAGCGCGGCGCGGGCTGGT    429
         |||  ||| || ||| || |||  || ||| ||||||  ||| || || || |||||
Db  657  GCTGGCCATCTGGGCGGGCTGTCCGCGCGGTGCGGCGCGCGGGTGCACACC          711
```

RESULT 12
US-09-576-290-103
; Sequence 103, Application US/09576290
; Patent No. 6794491
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Habata, Yugo
; APPLICANT: Kawamata, Yuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Kitada, Chieko
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/576,290
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/776,971
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/343371
; FILING DATE: 28-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7/343371
; FILING DATE: 12-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8/59419
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8/211805
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8/246573
; FILING DATE: 18-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026

Right column:

; REFERENCE/DOCKET NUMBER: 47176
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1331 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 118...1227
; OTHER INFORMATION:
US-09-576-290-103

Query Match   7.8%;   Score 79;   DB 4;   Length 1331;
Best Local Similarity   51.1%;   Pred. No. 9.9e-08;
Matches   212;   Conservative   0;   Mismatches 200;   Indels   3;   Gaps   1;

```
Qy   18  GGGCCTGGCGGGGGGGCTCACTGGTTGGGGCACGATGGGGGGTTCGGTGGTCCAACGGCTGGT   77
         |||  |||| ||||  |||  ||||  |||  ||| || ||| ||| || |  ||| || ||
Db  297  GGGGCTGATCGTGCTGCTGCTCTACAGCGTCGTGGGTGGTGGGGCTGGTGGGGGAACTGCT  356

Qy   78  GCTGCTCTGCCTGCTCTGCACACGGCGGGACATCCGCCGCCGCGGCGGCTCTTCACCCT   137
         |||||  ||||  |||   || ||   ||||   |||  |||  |||  ||| ||||
Db  357  GCTCGGTGCTCGGTGATCGGCGCGGGTGCGCCGCGGCTGACGGAACTTCCTCATCGG    416

Qy  138  GAACCTCACGTGCGGGGAACCTGCTGGTCACACATGCCGCTCACGCTGGCCGG        197
         ||||| ||  ||  |||  |||||||  |   ||| || |||||||||||||||
Db  417  CAACCTGGCCTTGTCCGACGTGCTCATGTGCACCGCCGCCTCAGCGCTGGCCTA       476

Qy  198  CGTCGTGGGGGCCAGCGGAGCCGGGG---GGCGACGCGCCTGTGCCGGCTGCCTTCCT   254
         |||||  ||| |||   ||  ||||   |||||| || |||||| ||  |||||||||
Db  477  TGCCTTCGAGCCACGGGGCGTGTTCGGCGGGGCGGGCCCTGTTGCCACCTGGTCTTCTTCCT  536

Qy  255  CGACACCTTCCTGGCTGCCAACTCCATGCTCAGCAATGGCCCGCGCTCAGCATCGACCGGCTG  314
         |||||||||||  ||||||||    |||   ||| ||| || || |||||||| || ||||||
Db  537  GCAGCCGGTCACCGTCTATGTGTTCACGGTTCACCTGGTCTTCTTCCT              596

Qy  315  GGTGGCCGTGGTCTTCCCGCTGAGCTACCGGGCCAAGATCGGCCTCCGAGCGCGGCGGCT   374
         ||||  ||| |||  ||| ||||||  ||  |||  |||  ||  | ||| || || |||
Db  597  CGTCCGTGCTGGTGCACCGGCCCATCTCGCTCGGGCCTCAGCGCCTACGGCTGT        656

Qy  375  CATGGTGGCCTACACGTGGGCTGCACGGCGCTCACCTTCCCAGCGCGGCGCGGGCTGGT    429
         |||  ||| || ||| || |||  || ||| ||||||  ||| || || || |||||
Db  657  GCTGGCCATCTGGGCGGGCTGTCCGCGCGGTGCGGCGCGCGGGTGCACACC           711
```

RESULT 13
US-08-540-650B-11
; Sequence 11, Application US/08540650B
; Patent No. 6399325
; GENERAL INFORMATION:
; APPLICANT: HINUMA, Shuji
; APPLICANT: FUJII, Ryo
; APPLICANT: FUKUSUMI, Shoji
; APPLICANT: OHTAKI, Tetsuya
; APPLICANT: HOSOYA, Masaki
; APPLICANT: OHGI, Kazuhiro
; APPLICANT: ONDA, Haruo
; TITLE OF INVENTION: GALANIN RECEPTOR PROTEIN, PRODUCTION AND USE THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: FastSEQ Version 1.5
SOFTWARE: DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/540,650B
FILING DATE: 11-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 7-134412
FILING DATE: 31-MAY-1995
APPLICATION NUMBER: 6-326610
FILING DATE: 28-DEC-1994
APPLICATION NUMBER: 6-247599
FILING DATE: 13-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: RESNICK, DAVID S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45901
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1882 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 722..1768
OTHER INFORMATION:
US-08-540-650B-11

Query Match   49.3%; Score 79; DB 3; Length 1882;
Best Local Similarity   49.3%; Pred. No. 1.1e-07;
Matches 264; Conservative   0; Mismatches 265; Indels   6; Gaps   2;

```
Qy     59 TGCTGTCCAACGCGCTGGTGTCGTCTGCTCTGCCTGCCTGCACAGC---GCGGACATCCGCCGCC 115
          |||| ||| |||| ||| ||  ||||   | |||  |||||| ||||||   |||||| ||||||
Db    864 TGCTGGGCAACAGCCTAGTGATCACCGTGCTGTGGCGGCGCAAGCGCGGGGCGAAGCGCGGGA 923

Qy    116 AGGGCGCGGCGGCCTCTTCACCCTGAACCTCAGTGCGGGAACCTGCTGCGCGTGGTCA 175
          ||||  || |   ||| ||  | |||   |    ||| | | || |  || | ||||
Db    924 GCACCACCAACCTGTTCATCCTCAACCTGAGCATCGCCGACCTGGCCTACTGCTCTTCT 983

Qy    176 ACATGCCGCTCACGCTGGCCGCGGGTCGTGGGGCCGCAGCGCGGACCGCGCCTGT 235
          |  | |||| ||| || || || |||||  |  ||  || | || || || || ||
Db    984 GCATCCCCTTCCAGGCACCGGTACGGCGTCCCCACCTGGGTGCTGGGCGCCTTCATCT 1043

Qy    296 CGCTCAGCAATCGACCGCCTGGGTGGGTGCGTGGTTCAGCTACCGGGGCCAAGAATGC 355
          || ||| || ||||| | || ||  | |  || | |||| | || ||  || || |
Db   1044 GCAAGTTCATCCACTACTTCTTCACCGTCCATCGTTGAGCAATCTTCACCTGGGCCG 1103

Qy    356 GCCTCCGCGACGCGGCGGCGCCTCATGGTGGCCTTACACGTGGCTTGCACGCGGCCTCACCTTCCAG 415
          ||||   |   || ||| || ||| ||| || ||| ||  | || |  ||  || ||||  |||||
Db   1104 CGATGTCCGTGGACGCCCTACGTGGCCCATCGTGCACTCGCCGCTCCTCCTCCTCAGG 1163

Qy   1164 TGTCCCGCAACGCGCGTGCTCGGCGTGGGCGTGGCGGTCCATCTGGGCGGCTGTCCATTGCCAT---GG 1220
          ||||
Db    356 ... 

Qy    416 CCGCGCCGGCGCTGGCCCTGTCTGGCTGGCCTTCCACCAGCTGTACGGCTTCGTGCACGGCCTGT 475
Db   1221 CCTCGCCCCGTGGCCTTACCACCAGGGCGCCTTCTTCCACCAGCAACCAGACCTTCT 1280

Qy    476 GCAGCCGGCGGCGGCCCAGACGAGCGCGCCTGGCCCTCTTCGCCGCTTCGCCGCCTTCCACGCGCTC 535
Db   1281 GCTCGGGAGCAGTGGGCCCGACCCCTACGTGGTTGTGTGCACCTTCGTCT 1340

Qy    536 TCAGCTTCTCTGCTTCTTTCTGTGCACGTACCTCAAGGTGTCTCAA 590
Db   1341 TCGGCTACCTGTCGGCTCCTGTCGCTCCTTCTCTGCTCATCTGCCTATGCCAAGGTCTCTTTAA 1395
```

Query Match   49.3%; Score 79; DB 3; Length 3083;
Best Local Similarity   49.3%; Pred. No. 1.2e-07;
Matches 264; Conservative   0; Mismatches 265; Indels   6; Gaps   2;

```
Qy     59 TGCTGTCCAACGCGCTGGTGTCGTCTGCTCTGCCTGCCTGCACAGC---GCGGACATCCGCCGCC 115
          |||| ||| |||| ||| ||  ||||   | |||  |||||| ||||||   |||||| ||||||
Db    929 TGCTGGGCAACAGCCTAGTGATCACCGTGCTGTGGCGGCGCAAGCGCGGGGCGAAGCGCGGGA 988

Qy    116 AGGGCGCGGCGGCCTCTTCACCCTGAACCTCAGTGCGGGAACCTGCTGCGCGTGGTCA 175
          ||||  || |   ||| ||  | |||   |    ||| | | || |  || | ||||
Db    989 GCACCACCAACCTGTTCATCCTCAACCTGAGCATCGCCGACCTGGCCTACTGCTCTTCT 1048

Qy    176 ACATGCCGCTCACGCTGGCCGCGGGTCGTGGGGCCGCAGCGCGGACCGCGCCTGT 235
          |  | |||| ||| || || || |||||  |  ||  || | || || || || ||
Db   1049 GCATCCCCTTCCAGGCACCGGTACGGCGTCCCCACCTGGGTGCTGGGCGCCTTCATCT 1108

Qy    236 GCCGCCTGGCGTGCCTTCCTCGACACCTTCCTGGCTGCCAACTCCATGCTCAGCAATGGCCG 295
          |
Db   1108
```

RESULT 15
US-08-985-090-3
Sequence 3, Application US/08985090
Patent No. 5885893
; GENERAL INFORMATION:
; APPLICANT: Andrew D.J. Goodearl
TITLE OF INVENTION: MUSCARINIC RECEPTORS AND USES THEREFOR
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,090
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jean M. Silveri
REGISTRATION NUMBER: 39,030
REFERENCE/DOCKET NUMBER: MNI-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1335 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1335
US-08-985-090-3

Query Match   52.7%;
Best Local Similarity
Matches 184;   Conservative   0; Mismatches 162;   Indels   3;   Gaps   1;

7.3%;   Score 73.8; DB 2;   Length 1335;
Pred. No. 1.3e-06;

Qy   8 CGTGGGACGCGGGGCCTGGCGGGGGGCTACTGGTGGGGCACGGATGGGGGTCGGCTGGTCCA   67
            ||| |||| || ||||| || ||| |||| ||| ||| |||||| || ||||||||
Db   95 CCTGGACGCGGGCGGTGCTGGGCCCGGGCCTCATGGGCCGCTGCTCATCGTTGGGCACGGGCA   154

Qy   68 ACGGCGCTGGGTGCTGCTGCCTTGCCTGCCACAGCGGCCGGACATCCGCCGCCGGGCCGCGGCC   127
            ||||||||| ||| ||||| |||| || ||| ||||| ||| |||||| |||||||| ||
Db   155 ACGCGCTGGGTCAGTCTGCCCTTCGTGTGCGGACTGTGAGCCTCCGCACCCAGGAACAACT---   211

Qy   128 TCTTCACCCTGAACCTCACGTGCGCAGGAAACCTTGCTGTTGCACCGTGGGTCAAACATTGCCGGCTCA   187
            |||||| ||||| ||||| ||| ||| |||| || |||| ||| ||||| | ||||| || |||
Db   212 TCTTCCTGCTCAAACCTGCGCCATCTCCGACTTCCTCGTCGGGCCCTTCTGCATCCCACACTGT   271

Qy   188 CGCTGGCCGCGGGGTCGTTGGGCGGCAGCGGCGGGCGGCGGACCGGCCGTCGTGCCCTGCCTG   247
            || | ||||| |||||| || || | ||||| |||||||||| |||||| || |||||||||
Db   272 ATGTACCCTACGTGCTTGACAGGCGGACCTTCGGCGCGGGGGCCTTCGCGAAGGCTGTGGGC   331

Qy   248 CCTTCCTCGACACCTTCCTCGGCTGCCAAACTCCATGCTTCAGCATGCGCGCGCTTCAGCATCG   307
            |||| || ||||| || || ||| | ||| || || ||| ||| |||| ||| ||
Db   332 TGGTAGTGGGACTACCCTGCTGTGCCACCTCCTCTGCCCTTCAACATCGTCGCTCATCAGCTTACG   391

Qy   308 ACCGGCTGGGGGTGGCGCCTGGGTCTTCCCGCGTGAGCTACCGGGCCAAGAGATCG   356
            |||||| | ||| || ||||| ||||| ||| |||| ||| |||||| || |||
Db   392 ACGGCTTCCTGTCGGTCACCCGAGCGCGGTCTCATACCGGGGGCCCAGCAGGG   440

Qy   1109 GCAAGTTCATCCACTACTTCTTCACCGTGTCCATGCTGGTGAGCAATCTTCACCCTGGGCCG   1168
            ||| ||| ||| ||| || |||| |||||| |||||
Db   1169 CGATGTCCGGTGGACCGGCTAGGTGGGCCATCGCGGGCGCTTCCTCCTCCCCTCAGGG   1228

Qy   1229 TGTCCGGCAACGGGCTGCTTGGGCGTGGGCGTTCCATCTCGTGTCCATTGCCAT---GG   1285
            || |||| ||| ||||||| ||| ||| ||||||||| || | || || |||||
Db   1286 CCTCGCCCGTGGCTACCACCAGGGCCTCTTCACCCGCCCAGCAACCAGACCTTCT   1345

Qy   416 CCGCGCGCTTGCCCTGTCTCTGGCGTCGGCGCGTTCACCAGCGCTGTACGCGCTGACGCGTGCACGGCTGT   475
            |||||| ||| ||||| || || |||||| | | | || ||| | ||| || || |||| | ||
Db   1346 GCTGGGAGCAGTGCCCGACCTCGCCACAAGAGAGGCCTACGTGGTGTGTGCACCTTCGTCT   1405

Qy   536 TCAGTTCCTGCTCTCCTTCGTCGTGTCTGCACGTACCTCAAGGTGCTCAA   590
            || ||||||||||| |||||| || ||||| ||| |||||||| ||||||
Db   1406 TCGGCTTACCTGCTGCCGCTCCTGCCTCATCCGCTTCTGCTATGCCAAGGTCCTTAA   1460

Search completed: July 5, 2005, 15:59:46
Job time : 212 secs

THIS PAGE BLANK (USPTO)